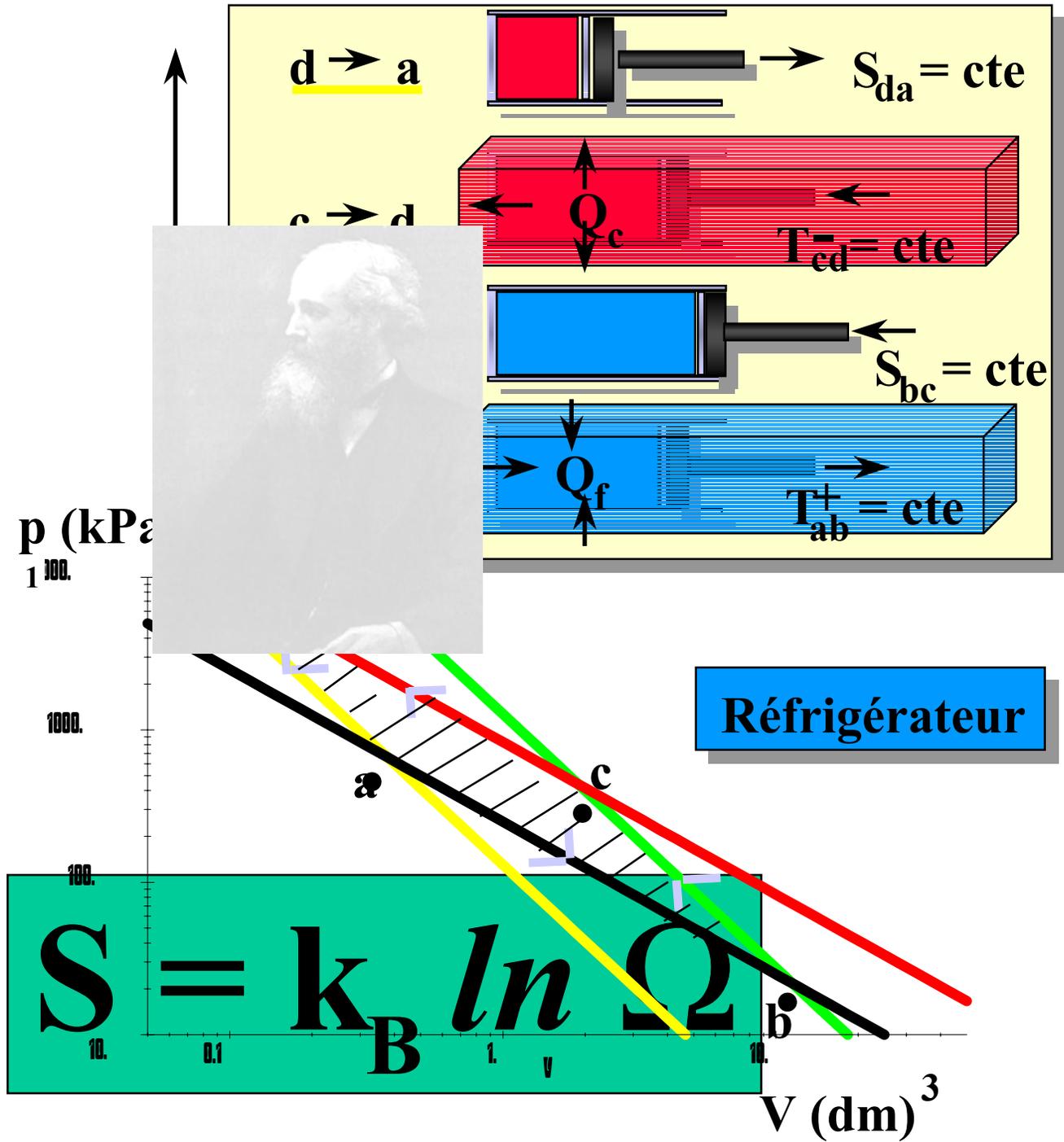


Physique statistique PHQ 340



Automne 2003, André-Marie Tremblay

Table des matières

0.1	Introduction générale	1
0.1.1	La place de la physique statistique dans le baccalauréat de physique et dans la physique	1
0.1.2	Bases de la physique	1
0.1.3	La physique statistique par rapport à la physique en général	3
0.1.4	* Une remarque “hors d’ordre” sur les unités	5
0.1.5	Les ”applications” de la physique statistique	6
0.1.6	Un bref aperçu du cours de physique statistique	6
0.1.7	Une brève histoire de la physique statistique et de la thermodynamique	7
1	Introduction aux méthodes statistiques	11
1.1	Un peu d’histoire ¹	11
1.2	Notions de base en théorie des probabilités ²	12
1.2.1	Ensembles et probabilités à priori ³	12
1.2.2	Probabilité d’événements plus complexes: ET et OU (lois de composition)	13
1.2.3	Valeur moyenne et écart type.	14
1.3	Analyse combinatoire	16
1.3.1	Permutations, arrangements et combinaisons	16
1.3.2	Binôme de Newton	19
1.4	Exemples de distributions de probabilité	20
1.4.1	La marche aléatoire: distribution de probabilité binomiale ⁴	20
1.4.2	Distribution de probabilité gaussienne	23
1.5	Changements de variables pour les distributions de probabilité continues	35
1.6	Discussion générale de la marche aléatoire et loi des grands nombres	40
1.6.1	Distribution de probabilité pour plusieurs variables ⁵	41
1.6.2	Discussion générale des valeurs moyennes pour la marche aléatoire ⁶	42
1.7	*Formule de Stirling	44
1.8	*Théorème de la limite centrale	45
1.9	*Chi carré et estimé de l’écart type	48
1.10	Résultats importants du chapitre 1	50
1.11	Problèmes pour le chapitre 1	53
1.11.1	Trouvez la bonne clé! ⁷	53
1.11.2	Jeux de cartes	53

¹Maury, p.150

²Reif, Sec. 1.1

³Feller tome 1 donne une introduction très détaillée à la notion de probabilité.

⁴Reif, Secs.1.1 et 1.2

⁵Reif, Sec.1.7

⁶Reif, Sec. 1.9

⁷Reif. Prob. 1.5

1.11.3	Le paradoxe des anniversaires!	53
1.11.4	Des juges impartiaux?	54
1.11.5	Le triangle de Pascal, d'intérêt historique	54
1.11.6	Mouvement Brownien	54
1.11.7	Moyenne et probabilités dans un jeu de hasard	55
1.11.8	Erreur sur la moyenne	55
1.11.9	Calculs avec la gaussienne	55
1.11.10	Quelques statistiques de radioactivité ⁸	56
1.11.11	Les erreurs de frappe ⁹	57
1.11.12	Élimination du bruit ¹⁰	58
1.11.13	Changements de variables ¹¹	58
1.11.14	Mesures expérimentales	58
1.11.15	Configurations de bosons	58
1.11.16	Configurations de fermions	59
1.11.17	Erreur quadratique moyenne	60
1.11.18	Pile ou face truqué	60
1.11.19	Sondages	60
2	Description statistique des systèmes physiques	61
2.1	Description statistique et postulat de base	62
2.1.1	Spécification de l'état <i>microscopique</i> d'un système ¹²	62
2.1.2	État <i>macroscopique</i>	63
2.1.3	Équilibre thermodynamique	64
2.1.4	Ensemble statistique microcanonique ¹³	65
2.1.5	Postulat de base	65
2.1.6	Calcul de probabilité ¹⁴	67
2.1.7	Dépendance en énergie de la densité d'états pour les systèmes macroscopiques ¹⁵	68
2.1.8	Dépendance en puissance de f des quantités physiques	71
2.2	Un peu d'histoire	72
2.3	Interactions entre systèmes macroscopiques	73
2.3.1	Interactions thermiques ¹⁶	74
2.3.2	Interactions mécaniques ¹⁷	75
2.3.3	Interactions quelconques, mécaniques et thermiques ¹⁸	78
2.3.4	Processus quasi-statique ¹⁹	80
2.3.5	Différentielles exactes et inexactes ²⁰	82
2.3.6	Fonction d'état et entropie	84
2.4	Équivalence entre travail et chaleur. Un peu d'histoire.	85
2.5	Résultats importants du chapitre	86
2.6	Problèmes pour le chapitre 2	87
2.6.1	Espace des phases ²¹	87
2.6.2	Interactions thermiques et mécaniques	88

⁸Reif, Prob. 1.12 et 1.9

⁹Reif, Prob. 1.11

¹⁰Reif, Prob. 1.21

¹¹Q2001

¹²Reif, Sec. 2.1

¹³Reif, Sec. 2.2

¹⁴Reif, Sec. 2.4

¹⁵Reif, Sec. 2.5

¹⁶Reif, Sec. 2.6

¹⁷Reif, Sec. 2.7

¹⁸Reif, Sec. 2.8

¹⁹Reif, Sec. 2.9, 2.10

²⁰Reif, Sec. 2.11

²¹Reif, Prob. 2.2

2.6.3	Différentielles exactes et inexactes ²²	88
2.6.4	Travail dans un processus quasi-statique ²³	88
2.6.5	Chaleur et travail dans un processus quasi-statique ²⁴	89
2.6.6	États accessibles pour un système de spins	89
2.6.7	Systèmes de spins en contact thermique	89
2.6.8	Fluctuations de densité dans un gaz ²⁵	90
2.6.9	État le plus probable pour deux gaz parfaits à l'équilibre	90
3	Lois de la thermodynamique	93
3.1	Irréversibilité et l'atteinte de l'équilibre	93
3.1.1	Relâchement des contraintes et augmentation du nombre d'états accessibles ²⁶	93
3.1.2	Processus réversibles et irréversibles ²⁷	95
3.2	Interactions thermiques et équilibre thermique	98
3.2.1	Distribution d'énergie entre systèmes à l'équilibre ²⁸	98
3.2.2	Valeur la plus probable, valeur moyenne et valeur à l'équilibre	101
3.2.3	Probabilité maximale, équilibre et température ²⁹	101
3.2.4	Probabilité et entropie ³⁰	102
3.2.5	Entropie totale et entropie de l'état macroscopique le plus probable ³¹	104
3.2.6	Changement d'énergie et d'entropie lors d'un contact thermique ³²	106
3.2.7	La mesure de la température: premier aperçu ³³	106
3.2.8	Propriétés de la température absolue. ³⁴	108
3.2.9	Réservoirs de chaleur ³⁵	111
3.2.10	Étroitesse de la distribution de probabilité (fluctuations) ³⁶	112
3.3	Interactions quelconques et équilibre général	114
3.3.1	Dépendance de la densité d'états sur les paramètres externes ³⁷	115
3.3.2	Conditions d'équilibre thermique et mécanique ³⁸	118
3.3.3	Équilibre de systèmes pouvant échanger des particules	119
3.3.4	Changements d'entropie et chaleur absorbée dans un processus quasi-statique infinitésimal ³⁹	120
3.3.5	Propriétés générales de l'entropie ⁴⁰	121
3.4	Résultats fondamentaux: lois de la thermodynamique et relations statistiques.	125
3.4.1	Lois de la thermodynamique ⁴¹	125
3.4.2	Relations statistiques ⁴²	127

²²Reif, Prob. 2.6

²³Reif, Prob. 2.8

²⁴Reif, Prob. 2.11

²⁵Reif, Berkely, Prob. 2.15

²⁶Reif, Sec. 3.1

²⁷Reif, Sec. 3.2

²⁸Reif, Sec. 3.3

²⁹Reif, Sec. 3.3

³⁰Reif, Sec. 3.3

³¹Reif, p.111

³²Reif, Sec. 3.4

³³Reif, Sec. 3.5

³⁴Reif, Sec. 3.5

³⁵Reif, Sec. 3.5

³⁶Reif, Sec. 3.7

³⁷Reif, Sec. 3.8

³⁸Reif, Sec. 3.9

³⁹Reif, Sec. 3.9, p.115

⁴⁰Reif, Sec. 3.10

⁴¹Reif, Sec. 3.11

⁴²Reif, Sec. 3.11

3.5	Calcul statistique de quantités thermodynamiques	127
3.5.1	Équation d'état pour les gaz parfaits ⁴³	127
3.5.2	Distributions de Fermi-Dirac, Bose-Einstein et Maxwell-Boltzmann.	129
3.6	Un peu d'histoire	132
3.6.1	Les fils de la révolution ⁴⁴	133
3.6.2	L'édifice thermodynamique ⁴⁵	133
3.6.3	Carnot ou Joule ⁴⁶	134
3.6.4	Clausius "y regarde de plus près" ⁴⁷	134
3.6.5	L'entropie ⁴⁸	135
3.6.6	Le genre de mouvement que nous appelons chaleur	135
3.7	Résultats importants du chapitre	136
3.8	Problèmes pour le chapitre 3	141
3.8.1	Irréversibilité ⁴⁹	141
3.8.2	Thermodynamique d'un système de spins ⁵⁰	142
3.8.3	Changement d'entropie du gaz parfait dans un processus irréversible.	143
3.8.4	Mélanges de gaz parfaits ⁵¹	144
3.8.5	Détente adiabatique d'un gaz parfait	144
3.8.6	Impuretés dans les semi-conducteurs	144
3.8.7	Entropie des supraconducteurs à haute température dans la limite $T \rightarrow \infty$	145
4	Paramètres thermodynamiques, machines thermiques.	147
4.1	Travail, chaleur et température	148
4.1.1	Travail et énergie interne ⁵²	148
4.1.2	Chaleur ⁵³	151
4.1.3	Température ⁵⁴	152
4.1.4	Capacité calorifique et chaleur spécifique ⁵⁵	159
4.2	Entropie	161
4.2.1	Mesures de l'entropie et transfert de chaleur réversible ⁵⁶	162
4.2.2	Conséquences de la définition absolue de l'entropie et sa pro- priété de maximum. ⁵⁷	164
4.3	Paramètres intensifs et extensifs	169
4.4	Les machines thermiques	170
4.4.1	Un peu d'histoire: les machines à feu.	171
4.4.2	Un peu d'histoire: Carnot ⁵⁸	176
4.4.3	Moteurs thermiques ⁵⁹	177
4.4.4	Le raisonnement de Carnot sur le mouvement perpétuel et les extensions de Kelvin et Clausius ⁶⁰	180

⁴³Reif, Sec. 3.12

⁴⁴Maury, p.67

⁴⁵Maury, p.79

⁴⁶Maury, p.83

⁴⁷Maury, p.83

⁴⁸Maury, p.85

⁴⁹Reif, Prob. 3.4

⁵⁰Reif, Probs. 2.4, et 3.2

⁵¹Reif, Prob. 3.5

⁵²Reif, Sec. 4.1

⁵³Reif, Sec. 4.2

⁵⁴Reif, Sec. 4.3

⁵⁵Reif, Sec. 4.4

⁵⁶Reif, Sec. 4.5

⁵⁷Reif, Sec. 4.6

⁵⁸Maury, p.75 à 77

⁵⁹Reif, Sec. 5.11

⁶⁰Maury, p.73-74

4.4.5	Le cycle de Carnot ⁶¹	183
4.4.6	Réfrigérateurs ⁶²	186
4.5	Résultats importants du chapitre	188
4.6	Problèmes pour le chapitre 4	191
4.6.1	Équilibre thermique et génération d'entropie ⁶³	191
4.6.2	Chaleur et entropie pour un gaz parfait ⁶⁴	191
4.6.3	Conséquences de la définition absolue de l'entropie ⁶⁵	192
4.6.4	Moteur à essence ⁶⁶	192
4.6.5	γ pour un gaz parfait	193
4.6.6	Expansion de l'univers ⁶⁷	194
4.6.7	Chauffage par contact	195
4.6.8	Fluctuations d'énergie	195
5	Applications de la thermodynamique	197
5.1	Équations de base et changements de variables	197
5.1.1	Lois de base ⁶⁸	198
5.1.2	Changements de variables	199
5.1.3	Deux identités pour les dérivées des fonctions de deux variables	202
5.2	Propriétés des gaz parfaits	203
5.2.1	Équation d'état et énergie interne ⁶⁹	204
5.2.2	Chaleurs spécifiques ⁷⁰	206
5.2.3	Calcul microscopique des chaleurs spécifiques ⁷¹	208
5.2.4	Entropie ⁷²	209
5.2.5	Dilatation ou compression adiabatique ⁷³	210
5.3	Potentiels thermodynamiques et relations de Maxwell	212
5.3.1	Énergie interne: variables naturelles S et V . ⁷⁴	213
5.3.2	Enthalpie: variables naturelles S et p . ⁷⁵	217
5.3.3	Énergie libre de Helmholtz: variables naturelles T et V . ⁷⁶	218
5.3.4	Énergie libre de Gibbs: variables naturelles T et p . ⁷⁷	219
5.3.5	Variables thermodynamiques conjuguées, mémorisation des relations de Maxwell ⁷⁸	220
5.3.6	Potentiels thermodynamiques et transformation de Legendre	222
5.3.7	*Propriétés de convexité des potentiels thermodynamiques, stabilité et principe de Le Chatelier	227
5.3.8	Autres propriétés des potentiels thermodynamiques, incluant le cas des processus irréversibles	230
5.4	Relations thermodynamiques pour une substance homogène quelconque	231
5.4.1	Entropie et énergie interne ⁷⁹	232

⁶¹Reif, Sec. 5.11

⁶²Reif, Sec. 5.12

⁶³Reif, Prob. 4.2

⁶⁴Reif, Prob. 4.3

⁶⁵Reif, Prob. 4.4

⁶⁶Reif, Prob. 5.25

⁶⁷Q 1058

⁶⁸Reif, p.152,153.

⁶⁹Reif, Sec. 5.1

⁷⁰Reif, Sec. 5.2

⁷¹Reif, p.157

⁷²Reif, Sec. 5.4

⁷³Reif, Sec. 5.3

⁷⁴Reif, p. 161

⁷⁵Reif, p.162

⁷⁶Reif, p.163

⁷⁷Reif, p.163-164

⁷⁸Reif, Sec. 5.6

⁷⁹Reif, Sec. 5.8

5.4.2	*Préliminaires mathématiques sur les jacobiens ⁸⁰	233
5.4.3	Chaleur spécifique à volume et à pression constante ⁸¹	234
5.5	Troisième loi et relations de Maxwell	237
5.6	Refroidissement des gaz: détente libre et étranglement	239
5.6.1	Détente libre ⁸²	239
5.6.2	Étranglement Joule-Thomson ⁸³	240
5.6.3	Origine microscopique du refroidissement ⁸⁴	242
5.7	Et le nombre de particules?	244
5.7.1	Entropie du gaz parfait et paradoxe de Gibbs	245
5.7.2	Potentiels thermodynamiques <i>ad nauseam</i>	246
5.7.3	Relation de Gibbs-Duhem, ou la magie de l'extensivité	247
5.8	Résultats importants du chapitre	249
5.9	Problèmes pour le chapitre 5	252
5.9.1	Loi de Stefan-Boltzmann et considérations thermodynamiques	252
5.9.2	Bande élastique ⁸⁵	253
5.9.3	Pression atmosphérique ⁸⁶	254
5.9.4	Vitesse du son ⁸⁷	254
5.9.5	Thermodynamique d'un paramagnétique ⁸⁸	255
5.9.6	Compresseur défectueux? ⁸⁹	255
5.9.7	Compressibilités ⁹⁰	256
5.9.8	L'équation de van der Waals ⁹¹	256
5.9.9	Chaleur spécifique à pression constante ⁹²	257
5.9.10	Liquides parfaits ⁹³	257
5.9.11	Entropie de mélange ⁹⁴	257
5.9.12	Le paradoxe du chauffage des vieilles maisons ⁹⁵	258
5.9.13	Enthalpie et entropie du gaz parfait	258
5.9.14	Irréversibilité, chaleur, entropie dans un processus Joule-Thomson	259
5.9.15	Trouver l'entropie à partir de l'énergie (exemple des gaz parfaits quantiques)	259
5.9.16	Cohérence thermodynamique pour les gaz parfaits quantiques d'électrons libres (potentiels de Massieu)	260
5.9.17	Relation de Gibbs-Duhem pour un système binaire et addition des compressibilités.	260
5.9.18	Enthalpie d'un gaz parfait	261

6 Méthodes de calcul: ensemble canonique 263

6.1	Ensembles représentatifs de diverses situations physiques	264
6.1.1	Ensemble microcanonique ⁹⁶	264

⁸⁰Landau et Lifshitz, *op. cit.* p.67

⁸¹Landau et Lifshitz, *op. cit.*p.67

⁸²Reif, Sec. 5.9

⁸³Reif, Sec. 5.10

⁸⁴Reif, p.182

⁸⁵Q1083

⁸⁶Reif, Prob. 5.7

⁸⁷Q 1020

⁸⁸Q 1078

⁸⁹Q 1014

⁹⁰LL69

⁹¹Q 1086, L1.15

⁹²Q1090

⁹³L1.9

⁹⁴Reif, Prob. 7.4

⁹⁵Q1074

⁹⁶Reif, Sec. 6.1

6.1.2	Ensemble canonique ⁹⁷	264
6.1.3	Exemples de calculs de probabilité dans l'ensemble canonique ⁹⁸	266
6.2	Calculs dans l'ensemble canonique et connexion avec la thermodynamique.	268
6.2.1	Calculs de valeurs moyennes dans l'ensemble canonique: la fonction de partition ⁹⁹	268
6.2.2	Fonction de partition pour N systèmes indépendants et extensivité.	270
6.2.3	Connexion entre la fonction de partition et la thermodynamique ¹⁰⁰	272
6.2.4	Lien entre les trucs de calcul dans l'ensemble canonique et les dérivées thermodynamiques	273
6.2.5	Équivalence et différences entre les ensembles canonique et microcanonique	274
6.3	Dérivation de l'ensemble canonique à partir de la méthode des multiplicateurs de Lagrange	275
6.4	Résultats importants du chapitre	280
6.5	Problèmes pour le chapitre 6	281
6.5.1	Énergie moyenne d'un système de spins ¹⁰¹	281
6.5.2	Modèle microscopique d'une bande élastique ¹⁰²	282
6.5.3	Impureté de spin $S = 1$	283
6.5.4	Anomalie de Schottky ¹⁰³	283
6.5.5	Polarisation ¹⁰⁴	283
6.5.6	Méthode des multiplicateurs de Lagrange ¹⁰⁵	284
6.5.7	Interprétation statistique de la désaimantation adiabatique	284
7	Applications simples de l'ensemble canonique	285
7.1	Approche générale	285
7.2	Le gaz parfait	288
7.2.1	Calcul de la fonction de partition ¹⁰⁶	289
7.2.2	Propriétés thermodynamiques: pression, entropie, énergie interne ¹⁰⁷	291
7.2.3	Paradoxe de Gibbs ¹⁰⁸	293
7.2.4	Validité de la limite classique ¹⁰⁹	294
7.3	Le théorème d'équipartition	297
7.3.1	Preuve du théorème d'équipartition ¹¹⁰	297
7.3.2	Applications simples du théorème d'équipartition ¹¹¹	298
7.3.3	Mouvement Brownien et une brève histoire de la vérification de la théorie atomique ¹¹²	302
7.4	Un cas quantique: modèle d'Einstein pour la chaleur spécifique	305

⁹⁷Reif, Sec. 6.2

⁹⁸Reif, Sec. 6.3

⁹⁹Reif, Sec. 6.5

¹⁰⁰Reif, Sec. 6.6

¹⁰¹Reif, Prob. 6.2

¹⁰²Q2011

¹⁰³Reif, prob. 6.6

¹⁰⁴Reif, Prob. 6.8

¹⁰⁵Reif, prob. 6.12

¹⁰⁶Reif, Sec. 7.2

¹⁰⁷Reif, Sec. 7.2, 7.3

¹⁰⁸Reif, Sec. 7.3

¹⁰⁹Reif, Sec. 7.4

¹¹⁰Reif, Sec. 7.5

¹¹¹Reif, Sec. 7.6

¹¹²Maury, p.161

7.5	Paramagnétisme	307
7.6	Théorie cinétique des gaz dilués à l'équilibre	311
7.6.1	Un peu d'histoire: Maxwell et la théorie cinétique des gaz. ¹¹³	312
7.6.2	Distribution des vitesses de Maxwell ¹¹⁴	313
7.6.3	Autres distributions de vitesses et valeurs moyennes ¹¹⁵ . . .	315
7.6.4	Nombre de molécules frappant une surface ¹¹⁶	318
7.6.5	Effusion ¹¹⁷	321
7.6.6	Pression et transfert de quantité de mouvement ¹¹⁸	322
7.7	Résultats importants du chapitre (moins théorie cinétique)	323
7.8	Problèmes pour le chapitre 7	325
7.8.1	Énergie moyenne associée à la rotation des molécules diatomiques:	325
7.8.2	Chaleur spécifique des molécules diatomiques:	326
8	Un bref résumé	329
8.1	L'objet de la thermodynamique	329
8.2	Lois de la thermodynamique	330
8.3	Conséquences macroscopiques des lois de la thermodynamique . . .	330
8.3.1	Machines thermiques	330
8.3.2	Existence des fonctions énergie et entropie pour le gaz parfait	332
8.4	*Potentiels thermodynamiques	335
8.5	Intermède probabiliste	337
8.5.1	Notions élémentaires	337
8.5.2	Analyse combinatoire	338
8.5.3	Exemples simples de distributions de probabilité	338
8.6	Physique statistique	339
8.6.1	Préambule : États microscopiques et macroscopiques	339
8.6.2	Ensemble microcanonique et équilibre thermodynamique . .	340
8.6.3	Postulat de base	340
8.6.4	Changements quasi-statiques et connexion avec la thermodynamique	341
8.6.5	Calcul des fluctuations	342
8.6.6	Calculs à partir de modèles microscopiques dans l'ensemble microcanonique	343
8.6.7	*Ensemble canonique	344
8.7	Les objectifs du cours et comment ils ont été atteints	347
A	Corrigé, PHQ 340, Phy. Stat. 01-10-19	349
A.1	Une marche aléatoire (<i>2.5 points</i>)	350
A.2	Ensemble microcanonique (<i>2 points</i>)	351
A.3	Machine cyclique dans le plan $T - S$. (<i>4 points</i>)	352
A.4	Raisonnement d'Einstein sur la nature corpusculaire de la lumière (<i>4.5 points</i>)	353
B	Corrigé, PHQ 340, Phy. Stat. 01-12-17	357
B.1	Entropie des supraconducteurs à haute température dans la limite $T \rightarrow \infty$ (<i>2.5 points</i>)	359
B.2	Propriétés d'un cycle de Carnot dans le plan $T - S$. (<i>3 points</i>) . .	360

¹¹³Maury, p.129-131, 133, 150-152

¹¹⁴Reif, Sec. 7.9

¹¹⁵Reif, Sec. 7.10

¹¹⁶Reif, Sec. 7.11

¹¹⁷Reif, Sec. 7.12

¹¹⁸Reif, Sec. 7.13

B.3	Conditions d'équilibre ¹¹⁹ (3.5 points)	360
B.4	Potentiels thermodynamiques et gaz parfaits (1 point)	362
B.5	Thermodynamique d'un gaz imparfait (1.5 points)	362
B.6	Population des niveaux de l'hydrogène ¹²⁰ (2 points)	363
B.7	Modèle d'Einstein (4.5 points)	363
C	Corrigé, PHQ 340, Phy. Stat. 02-10-15	365
C.1	Marche aléatoire, N petit, N grand (2.5 points)	366
C.2	Lois de la thermodynamique (2.5 points)	367
C.3	Compression adiabatique d'un gaz parfait (2.5 points)	367
C.4	Modèle d'Einstein pour un solide (4 points)	368
D	Corrigé, PHQ 340, Phy. Stat. 02-12-16	371
D.1	Un jeu de dés modifiés (3 points)	373
D.2	Coefficient de performance du réfrigérateur à absorption (3 points)	373
D.3	Le paradoxe du chauffage des vieilles maisons ¹²¹ (2 points)	375
D.4	Chaleur spécifique d'une molécule diatomique (1.5 point)	375
D.5	L'atome de Hubbard (3.5 points)	376
D.6	Compressibilités isotherme et isentropique (2.5 point)	377

¹¹⁹DGLR, II.11

¹²⁰DGLR, III.44

¹²¹Q1074

Liste des figures

0-1	Représentation schématique du lien entre les cours du baccalauréat de physique.	2
0-2	Tombe de Boltzmann à Vienne	9
1-1	Probabilité de gain pour un jeu de dés simple.	16
1-2	Arbre pour les permutations de N objets.	17
1-3	Distribution binomiale de valeur moyenne 4 et d'écart type 2.4. . .	22
1-4	Convergence de la binomiale vers la gaussienne. $N = 1, 2, 3$	24
1-5	Convergence de la binomiale vers la gaussienne. $N = 4, 5, 6$	25
1-6	Convergence de la binomiale vers la gaussienne. $N = 7, 10, 20$. . .	26
1-7	Convergence de la binomiale vers la gaussienne. $N = 30, 50, 100$. .	27
1-8	Distribution gaussienne de moyenne 4 et d'écart type 2.4.	34
1-9	Changement de variables $\phi = -2u$	36
1-10	Changement de variables $\phi = 2u^2$	37
1-11	Changement de variables, d'un angle à la projection du vecteur. . .	38
1-12	Illustration du théorème de la limite centrale.	40
1-13	Densité de probabilité pour la valeur de $\sum_{i=1}^N x_i^2/N$ lorsque les x_i^2 sont obtenus d'une Gaussienne centrée à zéro dont l'écart type est égal à un. La courbe la moins piquée est pour $N = 4$ et les autres sont pour $N = 10$, $N = 100$, jusqu'à $N = 250$ pour la plus piquée.	51
1-14	Probabilité pour que la valeur de $\sum_{i=1}^N x_i^2/N$ soit plus petite que la valeur indiquée sur l'abscisse lorsque les x_i^2 sont obtenus d'une Gaussienne centrée à zéro dont l'écart type est égal à un. La courbe la plus douce est pour $N = 4$ et les autres sont pour $N = 10$, $N = 100$, jusqu'à $N = 250$ pour la plus raide. (Le graphique a été obtenu par simple intégration par la règle du trapèze des résultats précédents plutôt qu'analytiquement),	51
1-15	Résultats de mesures primitives de distance.	59
2-1	Espace des phases, oscillateur harmonique.	63
2-2	Exemple d'ensemble microcanonique avec cinq spins.	66
2-3	Calcul de la probabilité d'occuper la position q pour un oscillateur harmonique.	67
2-4	Contact thermique	74
2-5	Interaction mécanique: Coupe d'un piston agissant sur un gaz. . .	76
2-6	Augmentation de l'énergie interne par un travail mécanique.	77
2-7	Conversion d'énergie électrique en énergie interne.	77
2-8	Piston séparant deux gaz pour illustrer différents types possibles d'interactions.	78
2-9	Quatre configurations pour deux gaz en contact.	79
2-10	Travail fait par la pression lors de l'expansion d'un volume.	81
2-11	Illustration de l'intégrale d'une différentielle inexacte le long de deux chemins différents.. . . .	82

2-12	Deux chemins différents correspondant à des travaux différents mais aux mêmes états initiaux et finaux.	84
3-1	L'ouverture d'un trou à travers d'une paroi permet au gaz en b) de se répartir également entre les deux récipients. Pour retourner à l'état original, il ne suffit pas de recomprimer le gaz comme en c) puisque ceci produit l'échauffement illustré d) ce qui implique qu'un contact thermique devient nécessaire pour retrouver la température initiale.	96
3-2	Deux gaz isolés de l'extérieur mais interagissant thermiquement . .	98
3-3	Thermomètre à gaz.	108
3-4	Équilibre mécanique et thermique.	114
3-5	Changement de la densité d'états sous l'influence d'un changement de contrainte. Notez que r numérote les états microscopiques, pas la distance.	116
3-6	Interaction "chimique".	119
3-7	Un niveau d'énergie en contact avec un réservoir de Fermions. . .	130
3-8	Susceptibilité magnétique d'un sel ionique (Loi de Curie).	142
4-1	Énergie interne et paramètre thermométrique.	149
4-2	Mesure de l'énergie interne à l'aide à la fois du travail mécanique et du travail électrique.	150
4-3	Mesure de la chaleur absorbée par la méthode des comparaisons. .	152
4-4	Vérification de la loi de Boyle-Mariotte.	155
4-5	Thermomètre à volume constant.	155
4-6	Thermomètre à gaz pour les rapports de température absolue. . . .	156
4-7	Calibration d'un thermomètre au point triple de l'eau.	157
4-8	Point triple de l'eau dans le diagramme de phase $p - T$	157
4-9	Piège magnéto-optique pour refroidir des atomes.	165
4-10	Bosons et fermions.	166
4-11	Surface de Fermi d'un gaz de fermions.	167
4-12	Séparation de deux gaz à l'aide de deux membranes semi-perméables.	168
4-13	Shéma de la machine de Papin.	172
4-14	Pompe de Newcomen.	173
4-15	Piston à double action de Watt.	174
4-16	Locomotive à vapeur: chaudière	174
4-17	Locomotive à vapeur: pistons	175
4-18	Locomotive à vapeur: bielle	175
4-19	Locomotive à vapeur: excentrique	175
4-20	Locomotive à vapeur: ensemble du mécanisme moteur.	175
4-21	Représentation schématique du travail et des échanges de chaleur dans une machine thermique (moteur).	179
4-22	Moteur à combustion interne quatre temps.	180
4-23	Détails d'un piston de moteur à combustion interne.	181
4-24	Raisonnement de Carnot sur l'existence d'une efficacité maximale. .	182
4-25	Cycle de Carnot.	184
4-26	Représentation schématique d'un réfrigérateur.	187
4-27	Cycle de Carnot pour la réfrigération.	189
4-28	Cycle de Otto dans le plan $p - V$	193
5-1	L'entropie comme fonction monotone croissante de l'énergie.	199
5-2	Contre exemple: Une fonction non-monotone.	200
5-3	L'entropie est une fonction concave.	201

5-4	Contre exemple: Une fonction monotone croissante mais qui n'est pas concave. On voit qu'il peut y avoir deux valeurs de l'entropie et de l'énergie pour une seule valeur de la température.	201
5-5	Construction graphique de la transformée de Legendre $z(y')$ de la fonction $y(x)$	223
5-6	Ambiguïté survenant lorsqu'on essaie d'utiliser la dérivée comme variable indépendante.	226
5-7	Détente Joule-Thomson	241
6-1	Polarisation d'un milieu contenant des impuretés.	284
7-1	Shéma pour le calcul de la probabilité pour frapper la surface. . . .	319
7-2	Chaleur spécifique de molécules diatomiques en fonction de T	327
7-3	Capacité calorifique C_V par mole n en unités de R la constante des gaz parfaits pour la molécule de H_2 . (Adapté de F.W. Sears et G.L. Salinger, Addison-Wesley, 1975).	327
8-1	Shéma pour le raisonnement de Carnot	331
8-2	Représentation schématique du travail et des échanges de chaleur dans une machine thermique (moteur).	333
8-3	Deux chemins pour changer l'état d'un gaz parfait.	334
8-4	Équilibre thermique dans l'ensemble microcanonique	340
A-1	Cycle dans le plan $T - S$	352
B-1	Cycle de Carnot dans le plan $T - S$	360
D-1	Capacité calorifique C_V par mole n en unités de R la constante des gaz parfaits pour la molécule de H_2 . (Adapté de F.W. Sears et G.L. Salinger, Addison-Wesley, 1975).	376

Liste des tableaux

0.1 Introduction générale

Dans cette brève introduction, nous présenterons

- La place de la physique statistique dans le baccalauréat de physique et dans la physique
- La place de la physique statistique dans la physique en général
- Un aperçu des "applications" de la physique statistique
- Un aperçu du cours de physique statistique
- Une brève histoire de la physique statistique et de la thermodynamique

0.1.1 La place de la physique statistique dans le baccalauréat de physique et dans la physique

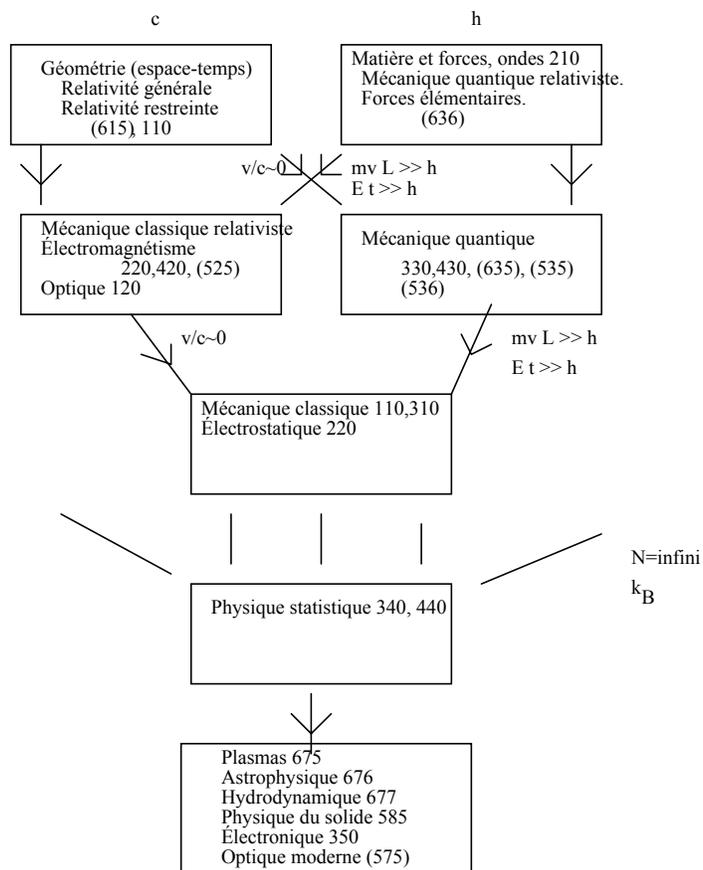
La mécanique statistique (ou physique statistique, c'est la même chose) constitue avec la mécanique quantique et la relativité l'un des piliers de la physique moderne. Elle a pour but d'expliquer le comportement de systèmes macroscopiques (incluant un très grand nombre de particules) à partir de leurs propriétés microscopiques. C'est de façon générale la mécanique quantique qui décrit les propriétés et l'évolution des systèmes physiques à l'échelle microscopique. La mécanique statistique est donc construite sur cette description quantique. Il est important de percevoir d'emblée la physique statistique comme une théorie fondamentale, et non pas comme une simple tentative de justifier *a posteriori* la thermodynamique. La démarche consiste donc à présenter la mécanique statistique élémentaire et à expliquer son articulation avec la thermodynamique et la théorie cinétique pour en dégager un point de vue unifié et moderne. La physique statistique permet ainsi de comprendre des phénomènes que les autres branches ne peuvent expliquer en raison d'une approche déterministe. Des applications concrètes peuvent être trouvées dans tous les domaines de la physique: physique des solides, électromagnétisme, astrophysique, cosmologie, superfluides, etc. Les deux cours de physique statistique sont obligatoires pour tous les étudiants.

La figure (0-1) donne une version schématique de la position de la physique statistique dans le baccalauréat de physique:

0.1.2 Bases de la physique

À la base de la physique, on retrouve donc

- Une théorie de l'espace-temps: la relativité (qui englobe la théorie de la gravitation lorsqu'on parle de relativité générale).



Entre parenthèses, cours optionnels

Remarques: 535 = compléments de méc. quantique

635 = mécanique quantique III

536 = physique atomique et moléculaire

Figure 0-1 Représentation schématique du lien entre les cours du baccalauréat de physique.

Une constante physique joue un rôle fondamental en relativité

$$\boxed{c} \tag{1}$$

la vitesse de la lumière. C'est en posant qu'elle est indépendante du système de référence que l'on dérive la relativité. Lorsque, dans un problème physique, les vitesses sont faibles par rapport à c , alors on retrouve l'invariance galiléenne.

- Une théorie du comportement de la matière dans cet espace-temps: la mécanique quantique. C'est la mécanique quantique qui explique le comportement des particules à l'échelle la plus microscopique. Évidemment, toute la structure et la richesse de l'univers vient du fait que ces particules interagissent entre elles pour former d'autres entités. La mécanique quantique décrit aussi les interactions entre les particules. Ces interactions entre les particules sont de trois types (mis à part la gravité): l'électromagnétisme, les forces faibles et les forces fortes. Ces deux dernières se manifestent surtout au niveau du noyau et des particules élémentaires. Aujourd'hui, on peut décrire ces trois interactions élémentaires dans un schéma unifié comme trois manifestations différentes de la même force.

En mécanique quantique on introduit une autre constante physique fondamentale qui a les unités de l'action (temps \times énergie, ou quantité de mouvement \times longueur), le "quantum d'action", la "constante de Planck"

$$\boxed{h} \tag{2}$$

Lorsqu'on s'intéresse à des problème où "l'action" est grande par rapport au quantum d'action, on retrouve la mécanique classique (relativiste ou non selon la valeur des vitesses par rapport à c). L'électromagnétisme (ou électrodynamique) peut être considéré dans la limite "classique" ou "quantique" mais demeure fondamentalement un phénomène relativiste. (Après tout, c est la vitesse de la lumière, une quantité fondamentalement électrodynamique!)

0.1.3 La physique statistique par rapport à la physique en général

Friction, chaleur.. pourquoi la physique statistique

À notre échelle, et dans presque tous les cas pratiques, la matière qui nous entoure se présente rarement une seule particule à la fois! Bien sur, si on s'intéresse au déplacement d'un rondelle de hockey, ou au déplacement de la lune ou d'un vaisseau spatial, on peut se contenter d'utiliser le théorème du centre de masse et la théorie des corps rigides en mécanique classique. Cette dernière vous sera présentée en détails dans le cours PHQ 310 à cette session-ci. Cependant, nous savons tous qu'il y a un type de phénomènes qui se présente à l'échelle macroscopique qui est à priori très différent de ce qu'on peut décrire par la mécanique classique, le phénomène de la *chaleur*. Celui-ci est intimement relié aussi à ce qu'on considère comme un embêtement dans les cours de mécanique classique: la friction. C'est en présence de friction que les lois de Newton perdent leur invariance sous inversion du temps. C'est en présence de friction et de "dissipation de chaleur" que l'on rencontre pour la première fois en mécanique un phénomène "irréversible". Qu'est ce que la chaleur, qu'est-ce que la température? Ce sont des concepts qui n'ont de sens que lorsqu'on considère un nombre très très grand de particules en interaction.

La base de la physique statistique, sa relation à la mécanique quantique et aux statistiques

La physique statistique se base sur les lois de la relativité et de la mécanique quantique décrites plus haut pour expliquer les comportements collectifs des assemblages de particules dans la limite où le nombre de ces particules tend vers l'infini. La physique statistique donc:

a) Utilise les lois fondamentales de la nature comme la mécanique quantique. En fait, comme nous le verrons, si on essaie de construire la physique statistique à partir des lois de la mécanique classique seulement, on arrive à de nombreuses contradictions. La physique statistique a été une des premières disciplines après la chimie à suggérer l'existence des atomes et des molécules. C'est aussi pour résoudre un problème de physique statistique, soit celui du corps noir, que Planck a postulé sa fameuse constante. Nous verrons que la mécanique quantique se manifeste à notre échelle macroscopique de façon frappante par le biais de la physique statistique.

b) Comme son nom l'indique, la physique statistique utilise des notions de probabilité, ou de statistiques, pour expliquer le comportement de grands nombres de particules. Plus les nombres sont grands, plus la théorie des probabilités a un sens. C'est un des nombreux résultats que nous discuterons au début du cours dans le contexte de notre étude des probabilités. Avec environ 10^{22} particules dans un gramme de matière, la physique statistique est certainement un des endroits où il est le plus sensé d'utiliser les statistiques! (Les sondages ne questionnent qu'un millier de personnes en général!). Dans la plupart des autres domaines où l'on utilise les statistiques, les nombres sont beaucoup plus petits et les méthodes utilisées sont plus sophistiquées que celles que nous discuterons ici. C'est tout un pan important du domaine des statistiques que nous devons laisser de côté, celui de "l'inférence statistique"

La physique statistique a sa constante fondamentale, ses notions fondamentales.

Il est très important de réaliser que la physique statistique nous forcera à introduire des concepts *qualitativement* nouveaux, propres au comportements collectifs des assemblages de particules. Par exemple, la physique statistique a elle aussi sa constante fondamentale

$$\boxed{k_B} \quad (3)$$

la constante de Boltzmann, qui permet de relier température et énergie par degré de liberté: $E \sim k_B T$ et de donner une signification "mécanique" (énergie) à la température. Cette constante fondamentale n'a aucun sens pour une seule particule. Elle n'est définie que dans la limite de la physique statistique, soit la limite des grands nombres de particules. La physique statistique a elle aussi ses lois fondamentales et ses concepts: température, entropie, enthalpie, énergie libre, pression, potentiel chimique, phase (eau-glace), fonction de partition.....

La physique statistique est à la base de la "thermodynamique" une théorie qui a été formulée *avant* la physique statistique sur des bases purement phénoménologiques. Cette "thermodynamique" décrit les phénomènes de la chaleur de façon *complètement générale*. On verra qu'à partir seulement de la conservation de l'énergie, de la structure quantique de la matière et d'un postulat probabiliste, on peut tirer des relations très générales entre des propriétés apparemment indépendantes de la substance: par exemple, la thermodynamique peut expliquer le fait que plus d'énergie soit nécessaire pour augmenter la température d'un corps d'un degré lorsque la pression est maintenue constante que lorsque le volume est maintenu constant. La thermodynamique se résume en trois *lois fondamentales*. Lisons-les une première fois ici, sous la forme la moins mathématique possible, même

si elles ne voudront rien dire avant que nous ayons introduit plusieurs notions et définitions

1. Le changement d'énergie dans un processus macroscopique est égal à la chaleur absorbée par le système, moins le travail mécanique fait par celui-ci.
2. Il est impossible de construire une machine qui transforme la chaleur en travail mécanique sans autre effet sur l'environnement (ou bien: Si deux systèmes à température différente sont mis en contact, la chaleur va toujours du système chaud au système froid et jamais l'inverse).
3. Il est impossible d'atteindre le zéro absolu.

La physique statistique va plus loin que la thermodynamique. Elle permet entre autres de calculer non seulement les valeurs moyennes d'observables, mais aussi les fluctuations. Lorsqu'elle est utilisée pour comprendre le comportement des gaz en terme des collisions élémentaires entre les particules, elle porte le nom de "théorie cinétique". Lorsqu'elle est utilisée pour comprendre des phénomènes quantiques, elle s'appelle mécanique statistique quantique etc...

0.1.4 * Une remarque "hors d'ordre" sur les unités

En mécanique, on définit trois unités arbitraires, à partir de standards. Dans les unités MKS par exemple, on définit un étalon (standard) de longueur, le mètre, un standard de temps, la seconde, et un standard de masse, le kilogramme. On pourrait tout aussi bien utiliser le pied, la seconde et la livre! On discute dans le cours de mécanique pourquoi ces trois définitions d'unités sont les seules nécessaires tant qu'on n'introduit pas l'électromagnétisme. (Il faut alors introduire l'ampère). La vitesse de la lumière se mesure dans les unités "mécaniques" que l'on a choisies, soit en mètres par seconde dans le système MKS. De ces mêmes unités, la constante de Planck se mesure en Joule-seconde ($kg \cdot m^2/s$). Ces deux quantités, vitesse de la lumière et constante de Planck, ne sont pas liées à l'introduction de nouvelles unités, simplement à la mesure de phénomènes physiques.¹²²

La constante de Boltzmann a un statut un peu différent de c et de \hbar . Si on avait décidé de mesurer la température en unités d'énergie et l'entropie dans des unités sans dimension, la constante de Boltzmann n'aurait pas été nécessaire. Elle est finalement là pour des raisons historiques, parce qu'on a défini une échelle de température avant de se rendre compte que température et énergie par degré de liberté sont en fait proportionnelles. Cependant, il est quand même bon de garder la distinction et de mesurer la température en Kelvin plutôt qu'en Joules car ceci permet d'insister sur le fait que la température est un concept *statistique* qui n'est pas défini pour une seule particule, alors que l'énergie elle l'est! Nous verrons que la température est simplement la dérivée de l'énergie moyenne par rapport à l'entropie, et cette dernière quantité est définitivement statistique!

¹²²Évidemment, on pourrait définir de nouvelles unités, dans lesquelles on définit un étalon de vitesse, un étalon d'action, en plus de la seconde. On peut aussi choisir ces étalons de telle sorte que $c = 1$ et $\hbar = 1$ dans ces nouvelles unités. Dans ces unités, les longueurs se mesurent en secondes- c , et les énergies en \hbar /seconde. Les facteurs de conversion sont que une seconde- $c = 2.9979250 \times 10^8 m$, alors que un \hbar /seconde = $1.054591847 \times 10^{-34} J$.

0.1.5 Les "applications" de la physique statistique

Un grand nombre de cours de la dernière année du baccalauréat de physique sont des "applications" de la physique statistique dans le sens où la physique statistique est importante pour comprendre plusieurs des phénomènes qui se produisent dans chaque domaine en particulier: Les plasmas, la physique du solide, l'hydrodynamique, l'astrophysique... sont des exemples de domaines basés sur la physique statistique. Sans parler de toute la chimie des réactions. Dans la session présente, on rencontre en électronique des applications de la physique statistique. Par exemple, pour comprendre la distribution en énergie des électrons dans un semiconducteur, il faut connaître la physique statistique. Vous rencontrerez aussi notre ami $k_B T$ dans la courbe courant-voltage d'une diode. On peut donc dire qu'on ne peut rien comprendre aux phénomènes physiques à la base de l'ensemble de l'électronique moderne, sans la physique statistique. Même en particules élémentaires, des analogies formelles entre la physique statistique et la théorie des champs permettent un échange continu entre ces disciplines: problème à N-corps, groupe de renormalisation, transition de phase...

0.1.6 Un bref aperçu du cours de physique statistique

Nous suivrons le développement du livre de F. Reif, "Fundamentals of Statistical and Thermal Physics". Ce même livre servira aussi au prochain trimestre pour PHQ-440. Les détails de la présentation en cours seront différents, mais les grandes lignes identiques.

Il faut d'abord introduire des notions élémentaires de probabilités et de statistiques avant de présenter au chapitre 2 la notion d'ensemble pour les systèmes mécaniques et la formulation statistique du problème mécanique. C'est là que nous introduirons le postulat statistique de base: Dans un système isolé, la probabilité d'un état *macroscopique* est proportionnelle au nombre d'états *microscopiques* correspondants à cet état macroscopique. Nous nous contentons ici des définitions intuitives de "microscopiques" et "macroscopiques", mais il est évident que nous devons définir ces notions plus en détails. En réfléchissant au type possible d'interactions entre systèmes macroscopiques, soit les interactions thermiques et mécaniques par exemple, nous pourrons "dériver" dans le chapitre 3 les lois de la thermodynamique à partir de notre postulat statistique de base. Nous décrirons ensuite dans le chapitre 4 comment, en pratique, on mesure des quantités thermodynamiques comme la température, la chaleur spécifique etc... et quels sont les principes de fonctionnement des machines thermiques (réfrigérateurs, machines à vapeur). Dans le chapitre 5, nous appliquerons ces idées aux gaz parfaits, et aux substances homogènes en général. Dans le chapitre 6, nous montrerons plus en détails comment faire un calcul à partir de la mécanique statistique à l'aide d'un concept bien commode, celui de la fonction de partition. Nous discuterons de sa relation avec la thermodynamique. Profitant de tous les outils que nous avons développé, au chapitre 7 nous pourrons faire les calculs qui permettent de comprendre les gaz parfaits (encore), le paramagnétisme, et la distribution de vitesse dans les gaz. Juste pour vous mettre l'eau à la bouche, je vous dévoile tout de suite que c'est au chapitre 8 tout de suite après Noël que vous verrez comment la mécanique statistique permet aussi de calculer les fluctuations en plus des valeurs moyennes ainsi que comment les choses se passent lorsque les substances peuvent échanger des particules. Cela nous permettra de comprendre l'équilibre entre phases, les réactions chimiques, en plus de nous permettre d'introduire la

notion de potentiel chimique et d'ensemble grand-canonique, indispensable pour les calculs dans le régime quantique.

0.1.7 Une brève histoire de la physique statistique et de la thermodynamique

Nous utiliserons le livre de J.-P. Maury et celui de Abraham Pais mentionnés dans la bibliographie comme guides dans les méandres de l'histoire de la physique statistique. Ces remarques historiques devraient permettre de faire des liens entre les différentes parties du cours, aider à mémoriser des concepts en leur associant une histoire, en plus de donner une perspective historique fascinante sur le développement d'une science. En effet, la physique statistique et la thermodynamique se sont développées sur une période d'un siècle environ. Et la présentation que nous en ferons en cours est une présentation "logique" qui est presque anti-historique. En d'autres mots, nous commençons par la perspective probabiliste, qui n'est venue qu'après la thermodynamique.

Alors qu'en électricité les découvertes fondamentales ont précédé les applications, en thermodynamique c'est tout le contraire. L'histoire de la thermodynamique c'est l'histoire des machines à vapeur qui furent à la base de la révolution industrielle. L'analyse formelle des machines à vapeur en termes fondamentaux n'est venue qu'en 1824 avec l'ingénieur français Sadi Carnot, dans un livre publié à compte d'auteur (600 exemplaires). Les industries n'ont pas attendu Carnot pour utiliser les machines à vapeur:¹²³

1698: Thomas Savery construit la première pompe à vapeur.

1712: Thomas Newcomen invente la pompe à feu.

1769: Joseph Cugnot construit le premier véhicule à vapeur, son fameux "fardier"

1780: James Watt construit le premier moteur à vapeur "à double effet".

1787: John Fitch construit le premier "vapeur".

1801: Richard Trevithick présente une voiture "sans chevaux" à vapeur.

1804: Richard Trevithick construit la première locomotive à vapeur.

1807: Robert Fulton, américain, construit le premier bateau à vapeur.

1880: Les voitures à vapeur commencent à devenir populaires.

Pendant un siècle, les locomotives à vapeur ont dominé le transport de marchandises et de passagers. En passant, il ne faut pas regarder les machines à vapeur avec dédain: même l'énergie nucléaire utilise la vapeur pour actionner les turbines électriques.

Du côté théorique, ce que Carnot a découvert en 1824 c'est la notion d'entropie, un concept assez abstrait. Cette découverte a eu lieu bien avant qu'on ait compris la relation entre chaleur et énergie! Cette dernière notion vous apparaîtra beaucoup plus naturelle que celle d'entropie. En fait elle est très facile à comprendre lorsqu'on sait que la matière est faite d'atomes. Et c'est l'ignorance de la validité de l'hypothèse atomique qui a sans doute rendu si difficile la découverte de l'équivalence entre énergie et chaleur. C'est peut-être une des raisons qui a poussé Feynman à faire l'énoncé suivant: *Si un cataclysme détruisait toute la connaissance scientifique et que seulement un énoncé était légué aux générations suivantes de créatures, quelle serait la phrase qui contiendrait le plus d'information dans le moins de mots. Je crois que ce serait la suivante: Toutes les choses sont faites d'atomes, c'est-à-dire de petites particules qui se déplacent, s'attirant lorsqu'elles ne sont pas trop loin, et se repoussant lorsqu'elles sont proches* (Traduction libre de *Feynman Lecture on Physics, Vol.1, p.I-2*).

¹²³Tiré de "L'âge de la vapeur", par Jonathan Rutland (Biblio. municipale, 621.2 R978a)

Toujours est-il que des anglais, le comte Rumford en 1798 et Davy en 1799 avaient suggéré cette équivalence entre chaleur et énergie. Le physicien allemand R.J. Mayer avait même énoncé explicitement cette équivalence. C'est cependant à Joule (1843-1849) que l'on doit les expériences qui *prouvent* cette équivalence. Le pauvre Joule, un amateur n'ayant pas besoin de travailler pour vivre, ne parvenait pas à faire publier ses travaux tellement cette idée apparaissait révolutionnaire à l'époque. Les *Proceedings of the Royal Society* ont refusé ses articles. *Philosophical Magazine* a fini par se laisser convaincre! En Allemagne, Helmholtz rencontrait les mêmes difficultés avec les mêmes idées!

Ce sont finalement Clausius et Kelvin aux environs de 1850 qui ont formulé la thermodynamique telle que nous la connaissons. Et c'est à Gibbs qu'on doit la première mention du terme "mécanique statistique" et les principes de bases du sujet. Nous utilisons "physique statistique" ou "mécanique statistique" de façon interchangeable dans ce cours. Les travaux de Gibbs dans ce domaine parurent en 1876-1878. Et tout ça s'est passé non loin d'ici. Il vous suffit de prendre l'autoroute plein sud et de conduire pendant cinq heures pour arriver à Yale. Gibbs a été le seul physicien d'envergure internationale en Amérique du Nord au XIXème siècle. Il travaillait sans salaire à Yale car il avait sa propre fortune. Il a publié ses premiers travaux dans les *Transactions of the Connecticut Academy of Arts and Sciences!* Pas très connu. On ne peut pas passer sous silence la contribution d'Einstein à la physique statistique¹²⁴. Deux grandes questions le préoccupait: La réalité des molécules et les fondements moléculaires de la physique statistique. En 1905, la fameuse année de la découverte de la relativité, Einstein a trouvé trois façons de compter les molécules, toutes donnant la même réponse. En mars, il compta les molécules (donna un estimé du nombre d'Avogadro) dans son article sur les quanta de lumière, en avril il le fit en utilisant les propriétés d'écoulement de molécules de sucre dans l'eau, et en mai il le fit en même temps qu'il expliquait la théorie du mouvement Brownien, sans doute une de ses plus célèbres contributions à la physique statistique. De ses contributions aux fondements de la physique statistique, on peut retenir¹²⁵ sa contribution profonde à la compréhension de ce qu'il a appelé "*le principe de Boltzmann*" $S = k_B \ln \Omega$, qu'il a plutôt utilisé sous la forme $P = \Omega = \exp(S/k_B)$ où il permet de calculer la probabilité d'un état à partir d'une quantité macroscopique, l'entropie. Einstein s'est intéressé à la physique statistique de façon active entre 1902 et 1925. Un des phénomènes qu'il a découvert, la condensation de Bose-Einstein, a été observée il y a longtemps dans l'hélium liquide, mais ce n'est qu'en 1995 qu'on a pu l'observer dans des gaz. C'est donc un sujet brûlant d'actualité, une découverte qui a été rendue possible grâce au refroidissement par laser!

Revenant un peu en arrière dans le temps, ce sont Clausius, Maxwell et Boltzmann qui ont travaillé très tôt sur les conséquences de *l'hypothèse atomique*. Dès 1859, Maxwell découvrit la loi de distribution des vitesses d'un gaz et Boltzmann en 1872 donna une analyse microscopique de l'irréversibilité et de l'atteinte de l'équilibre. La loi

$$S = k_B \ln \Omega \tag{4}$$

reliant entropie et nombre d'états microscopique accessibles apparaît sur sa tombe à Vienne. Nous reviendrons souvent sur cette loi fondamentale. On peut noter, curieusement que c'est plutôt Planck qui a introduit la constante de Boltzmann k_B .¹²⁶

¹²⁴ Pais, p.18

¹²⁵ Pais

¹²⁶ Pais, p.60

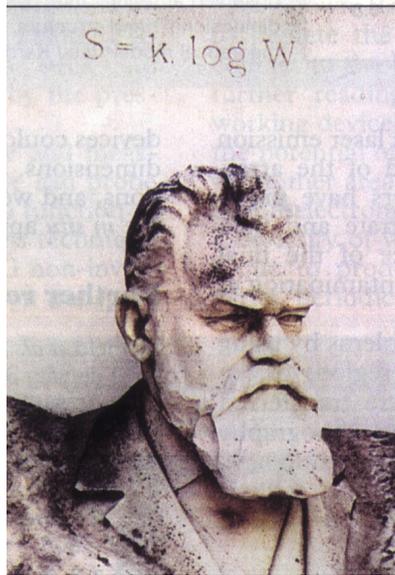


Figure 0-2 Tombe de Boltzmann à Vienne

1. INTRODUCTION AUX MÉTHODES STATISTIQUES

C'est dans ce chapitre que nous nous familiariserons avec les méthodes probabilistes et statistiques d'un point de vue très général. À la fin de ce chapitre, vous devriez pouvoir calculer vos chances de gagner à certains jeux de hasard en utilisant la notion de probabilité et les règles pour combiner les probabilités (et,ou). L'analyse combinatoire vous aura donné les outils pour compter rapidement toutes les façons, par exemple d'arranger N objets en différents groupes. Vous pourrez calculer les propriétés de base de la marche aléatoire, un modèle qui a donné la clé de la compréhension du mouvement Brownien. La marche aléatoire est bien décrite par la distribution de probabilité dite "binomiale". Dans la limite où le nombre de pas devient très grand, celle-ci devient la distribution de probabilité gaussienne. Dans une autre limite, qui correspond plus à celle qu'on rencontre dans les problèmes de radioactivité, cette distribution de probabilité devient la distribution de probabilité de Poisson. Vous pourrez donc par exemple calculer les probabilités d'obtenir par exemple trois désintégrations dans une période de cinq secondes d'une substance radioactive! Nous reviendrons à la fin sur une discussion plus générale de la marche aléatoire pour que celle-ci n'aie plus de secrets pour vous! Nous comprendrons même peut-être le théorème de la limite centrale, la clé de beaucoup de problèmes de statistiques! Ce théorème nous montre que la gaussienne a beaucoup plus d'importance que seulement comme cas limite de la binomiale!

1.1 Un peu d'histoire¹

C'est dans une correspondance de 1654 entre Pascal (1623-1662) et Fermat (1601-1665) qu'on retrouve les premiers écrits sur la théorie des probabilités et l'analyse combinatoire. La discussion portait sur un problème posé par le Chevalier de Méré (1610-1685). Pascal et Fermat se sont mis d'accord sur la solution de ce problème, bien qu'ils y soit parvenu par des méthodes différentes. Pascal expliqua sa méthode dans une série de petits articles qui ne furent publiés qu'en 1665, après sa mort. Sa méthode de solution passait par le "triangle de Pascal" que nous reverrons plus loin.

Presque deux cents ans plus tard, à l'été 1850, on retrouve le jeune Maxwell à dix-neuf ans. C'est probablement à cet âge qu'il a commencé à s'intéresser à la statistique. Il a lu un long article de J. Herschel dans *Edinburgh Review* qui relatait les travaux de Quételet, astronome et mathématicien belge, qui faisait valoir l'usage des méthodes statistiques en sciences humaines. En fait, le premier usage connu du mot "statistique", dérivé de l'Allemand, avait simplement signifié "Science de l'État", et au cours du *XVIII^e* siècle il décrivait l'étude des constitutions, des ressources nationales et des politiques des états. Drôle de lecture d'été pour ce jeune Maxwell. Qui devient quand même le premier, inspiré par Clausius

¹Maury, p.150

en 1858, à placer les statistiques et les probabilités à la base de la théorie des gaz, la cinétique. Un pas de géant était fait. À cette époque il faut se souvenir que l'idée de molécules était considérée comme très spéculative.

1.2 Notions de base en théorie des probabilités²

La théorie des probabilités est un vaste domaine aux applications innombrables. Nous ne ferons que l'effleurer. En particulier, nous ne verrons à peu près pas les notions d'échantillonnage et d'inférence statistique. C'est le fait qu'en physique on travaille presque toujours avec un nombre très grand de particules qui fait que les notions de probabilité que nous devons utiliser sont élémentaires. L'ouvrage en deux tomes de W. Feller sur les probabilités est une des meilleures références sur les probabilités du point de vue du physicien.

1.2.1 Ensembles et probabilités à priori³

Pile ou face. Un jeu bien connu. Prédire si une pièce de monnaie tombera définitivement du côté pile ou du côté face dans une expérience donnée est un problème d'une trop grande complexité pour espérer en trouver la solution. Nous pouvons par contre facilement nous convaincre que le côté pile et le côté face étant équivalents, en répétant l'expérience un grand nombre de fois la pièce tombera autant de fois du côté pile que du côté face. À partir de cet exemple simple on peut comprendre empiriquement les notions suivantes:

Ensemble: Imaginons N tirages à pile ou face. Prenons $N \rightarrow \infty$. L'ensemble de ces tirages à pile ou face est celui qui nous intéresse dans cet exemple. (Plutôt que de référer à "l'ensemble", on réfère parfois à "L'Univers des possibles")

Probabilités: La probabilité est définie par rapport à l'ensemble. La probabilité d'obtenir pile dans notre ensemble est donnée par la fraction des expériences qui donnent pile. On s'attend à ce que sur N tirages à pile ou face, $N/2$ fois on obtienne pile. Ainsi, la probabilité d'obtenir pile est $(N/2)/N = 1/2$.

De façon plus abstraite, on définit un ensemble "d'événements". Chacun des événements de l'ensemble est étiqueté par i . Cet ensemble constitue un espace sur lequel on définit une *probabilité à priori*. En d'autres mots, à chacun des différents "événements" de notre ensemble, on associe un nombre appelé $P(i)$ compris dans l'intervalle $[0, 1]$. En accord avec la notion intuitive de probabilité, on a que

$$\sum_i P(i) = 1 \quad (1.1)$$

La somme est sur tous les événements différents possibles. On appelle souvent l'équation ci-dessus la *condition de normalisation*.

Exemple 1 Dans le jeu de pile ou face les probabilités à priori sont les suivantes: $P(\text{pile}) = 1/2, P(\text{face}) = 1/2$ et la condition de normalisation est évidemment satisfaite.

²Reif, Sec. 1.1

³Feller tome 1 donne une introduction très détaillée à la notion de probabilité.

En d'autres mots, les probabilités à priori pour chacun des événements de l'ensemble sont les hypothèses de départ. On peut avoir une idée "expérimentale" de ce qu'est la probabilité d'un événement dans l'ensemble, mais ultimement en théorie des probabilités les probabilités à priori font partie des hypothèses de départ.

La notion de probabilité devrait vous être très claire intuitivement. Toute la théorie des probabilités consiste ensuite à prédire la probabilité de résultats d'expériences plus complexes à partir des probabilités à priori. Par exemple: Quelle est la probabilité d'obtenir deux pile et une face si on fait trois tirages à pile ou face consécutifs? À chaque fois par probabilité d'un événement (deux pile, une face) on veut dire: Si je répète l'expérience un grand nombre de fois, quelle fraction de fois vais-je obtenir l'événement en question. Par définition donc, la probabilité d'un événement est toujours comprise entre 0 et 1.

Remarque 1 *La probabilité dépend de l'ensemble de départ, en d'autres mots de l'information déjà connue. Par exemple, la probabilité qu'une graine de tulipe donne une tulipe rouge sera différente selon que cette graine est tirée a) d'un ensemble de graines donnant des tulipes rouges ou jaunes b) d'un ensemble de graines de tulipes de couleurs arbitraires.*

Remarque 2 *Voici un autre exemple illustrant que la probabilité est une notion qui dépend de l'information disponible. Durant le célèbre procès de O.J. Simpson dans les années 90, l'avocat défendant O.J. Simpson a répété à plusieurs reprises que le fait que O.J. Simpson battait sa femme n'était pas pertinent pour la preuve de l'accusation puisque seulement une femme battue sur mille est tuée par son mari. Évidemment, dans ce cas-ci on savait que sa femme avait été tuée, la statistique pertinente dans ce cas était plutôt que 80% des femmes battues qui ont été tuées l'ont été par leur mari! (L'autre morale de cette histoire est qu'il faut toujours se demander qui utilise quelles statistiques et pourquoi.)*

1.2.2 Probabilité d'événements plus complexes: ET et OU (lois de composition)

Cette section illustre comment on utilise les probabilités à priori de nos événements de départ pour trouver la probabilité d'événements plus complexes. Commençons par un exercice simple

Exercice 1.2.1 *Supposons que j'ai un dé à six faces. Je définis la probabilité à priori d'obtenir n'importe quelle face comme étant $1/6$ dans un ensemble d'expériences ou je lance un dé. En d'autres mots, le dé n'est pas truqué! Quelle est la probabilité d'obtenir un trois ou un cinq au premier lancer? ?*

Pour résoudre cet exercice, intuitivement on fait appel à la loi suivante.

- Si deux événements sont *complémentaires*, c'est-à-dire s'excluent mutuellement, la probabilité d'observer l'un OU l'autre est égale à la somme des probabilités de chacun des événements pris séparément.

Dans l'exemple des dés, les événements sont complémentaires car si j'obtiens trois je n'obtiens pas cinq et vice-versa. Essayons de trouver le résultat à l'aide de la notion intuitive de probabilité. Si je regarde les \mathbb{N} éléments de l'ensemble, j'obtiendrai $\mathbb{N}/6$ fois le nombre trois et $\mathbb{N}/6$ fois le nombre cinq. J'obtiendrai donc trois *ou* cinq, $2\mathbb{N}/6$ fois. La probabilité d'obtenir trois *ou* cinq est calculée en divisant par le nombre total d'éléments dans l'ensemble, soit \mathbb{N} , donc $(2\mathbb{N}/6)/\mathbb{N} =$

1/3. En d'autres mots, la probabilité est égale à $(\mathbb{N}/6) / \mathbb{N} + (\mathbb{N}/6) / \mathbb{N} = 1/6 + 1/6 = 1/3$. Finalement donc, il a suffi d'additionner les probabilités.

Exercice 1.2.2 Avec le même dé, quelle est la probabilité d'obtenir un trois au premier lancer et un cinq au deuxième lancer

Pour solutionner intuitivement ce problème, retournons à notre ensemble de \mathbb{N} éléments. Pour chacun des éléments correspondant au premier lancer, il y en a \mathbb{N} autres au deuxième, donc en tout \mathbb{N}^2 éléments dans notre nouvel ensemble. Le nombre de fois qu'un trois arrive au premier lancer est égal à $\mathbb{N}/6$. Pour chacun de ces heureux événements, je trouverai $\mathbb{N}/6$ fois un cinq dans le deuxième lancer. On retrouve donc un trois au premier lancer et un cinq au deuxième $\mathbb{N}/6 \times \mathbb{N}/6$ fois, ce qui correspond à une probabilité de $(\mathbb{N}/6 \times \mathbb{N}/6) / \mathbb{N}^2 = 1/6 \times 1/6 = 1/36$. En d'autres mots, la probabilité pour obtenir un trois au premier lancer et un cinq au deuxième est égal au produit des probabilités de chacun des événements pris séparément. On peut formaliser ce résultat de la façon suivante:

- Si deux événements sont *indépendants statistiquement*, en d'autres mots, si la probabilité d'en obtenir un n'influence pas la probabilité d'en obtenir un autre, la probabilité d'obtenir un ET l'autre est égal au produit des probabilités d'obtenir chacun séparément.

Remarque 3 Nous venons de définir l'indépendance statistique. Si deux événements ne sont pas indépendants statistiquement, il faut explicitement spécifier la probabilité conjointe, c'est-à-dire la probabilité d'obtenir un et l'autre.

Finalement, voici un exemple où il faut utiliser les deux lois ci-dessus. Supposons qu'on lance deux dés simultanément (un bleu et un rouge) et qu'on veuille savoir la probabilité d'obtenir trois et cinq indépendamment de la couleur du dé ayant le trois ou le cinq. Cette probabilité est de $1/18$. En effet, on sait déjà que la probabilité d'avoir trois sur le dé bleu et cinq sur le dé rouge est $1/36$. De même, la probabilité d'avoir cinq sur le bleu et trois sur le rouge est de $1/36$. La probabilité d'obtenir un trois et un cinq, indépendamment de la couleur du dé, est la somme des probabilités de ces deux événements exclusifs.

1.2.3 Valeur moyenne et écart type.

Dans cette sous-section, nous discuterons de deux quantités numériques souvent rencontrées dans le calcul des probabilités, la valeur moyenne, et l'écart type.

Si à chaque événement je peux associer un nombre, alors je peux obtenir la valeur moyenne de ce nombre. Pour plaire à Loto-Québec, supposons que si j'obtiens pile, je gagne 1.00\$ et si j'obtiens face, je perds 1.10\$. Quel est le gain moyen si je joue un très grand nombre de fois? Je perds en moyenne 0.05\$ par coup joué. En effet, si je joue \mathbb{N} fois, je gagnerai 1.00\$ un nombre de fois égal à $\mathbb{N}/2$ et je perdrai 1.10\$ le même nombre de fois, soit un gain de $(1.00\$ - 1.10\$) \times \mathbb{N}/2$ sur un nombre total de \mathbb{N} parties. Un gain de $-0.05\$$ par partie!

Remarque 4 Nous venons de faire référence au nombre \mathbb{N} pour la dernière fois. L'ensemble sera défini en même temps que les probabilités a priori et les deux petites règles du ET et du OU ci-dessus suffiront pour faire tous les calculs.

De façon plus abstraite, appelons $P(i)$ la probabilité d'obtenir le résultat i . En supposant que i ne peut prendre que les valeurs 1 à N , alors

$$\boxed{\sum_{i=1}^N P(i) = 1.} \tag{1.2}$$

En d'autres mots, $P(i)$, comme toute bonne distribution de probabilité, est "normalisée". La probabilité d'obtenir n'importe lequel des résultats possibles est évidemment égale à l'unité.

Soit maintenant $u(i)$ une valeur numérique associée au résultat i , (par exemple le nombre de dollars associé au résultat pile ou face) alors on définit la valeur moyenne \bar{u} ou $\langle u \rangle$ de u par

$$\bar{u} = \langle u \rangle = \sum_{i=1}^N P(i) u(i) \quad (1.3)$$

où N est le nombre total de résultats possibles. Par exemple, $N = 2$ pour pile ou face et $N = 6$ pour les dés. Les deux notations, \bar{u} ou $\langle u \rangle$ sont souvent utilisées pour noter une valeur moyenne. Une autre façon de dire la même chose est de dire que les symboles $\langle \dots \rangle$, ou $\overline{\dots}$ veulent dire de faire la moyenne de la quantité \dots de la façon suivante: $\sum_{i=1}^N P(i) \dots$.

Exemple 2 Supposons maintenant que notre joueur s'intéresse au jeu de dés suivant.

Nombre sur le dé	1	2	3	4	5	6	
Gain	-3\$	-2\$	-1\$	0\$	1\$	2\$	Dans ce cas, on peut

utiliser le nombre écrit sur le dé pour caractériser le résultat de l'expérience. Notons ce résultat par i . Si le dé n'est pas truqué, la probabilité d'obtenir i est donnée par $P(i) = 1/6$. On voit d'après le tableau que le gain $u(i)$ pour l'événement i peut être décrit par la formule suivante: $u(i) = i - 4$ dollars. Le gain moyen est donc

$$\bar{u} = \langle u \rangle = \sum_{i=1}^6 \frac{i-4}{6} \$ = \frac{(21 - 6 \times 4)}{6} \$ = -0.50\$ \quad (1.4)$$

Supposons que notre joueur s'intéresse aux "fluctuations" de son gain moyen. Le mot fluctuations est plutôt vague mais son sens intuitif est assez clair: Parfois on gagne, parfois on perd, est-ce que le montant gagné ou perdu à chaque fois est très différent ou non de la moyenne. Il y a plusieurs façons de répondre à cette question, et si on connaît la loi de probabilité, en fait on connaît tout ce qu'il est possible de connaître sur ce jeu. Néanmoins, on voudrait avoir un seul nombre qui caractériserait les fluctuations. Par convention, on se sert de l'écart type, aussi connu sous le nom de dispersion. Cette quantité est définie de la façon suivante:

$$\sigma \equiv \sqrt{\langle (u - \langle u \rangle)^2 \rangle} \quad (1.5)$$

Encore une fois, les crochets $\langle \rangle$ veulent dire "valeur moyenne". Dans l'exemple qui nous préoccupe

$$\langle (u - \langle u \rangle)^2 \rangle = \sum_{i=1}^6 \frac{1}{6} (i - 4 + 0.50)^2 \$^2 = 2.92\$^2 \quad (1.6)$$

donc, l'écart type σ est égal à $\sqrt{2.92\$^2} = 1.71\$$. Son gain peut donc fluctuer de presque 2\$ autour de la valeur moyenne, ce qui semble bien caractériser le jeu décrit ci-haut.

Définition 3 La quantité σ^2 s'appelle l'écart quadratique moyen ou variance.

On peut aussi utiliser le gain lui-même, u comme variable aléatoire. On a alors, $P(u) = 1/6$ pour toute valeur de u . La distribution de probabilité, sa valeur moyenne et son écart type se représentent graphiquement comme à la figure (1-1). La flèche verticale indique la valeur moyenne et l'horizontale l'écart-type.

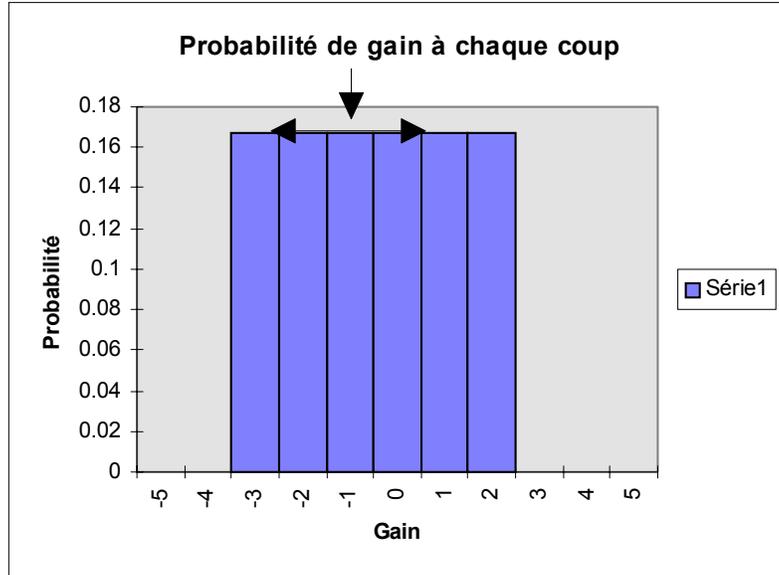


Figure 1-1 Probabilité de gain pour un jeu de dés simple.

Définition 4 On appelle variable aléatoire, une variable dont la valeur n'est connue qu'avec une certaine probabilité.

Remarque 5 On définit $\langle u^n \rangle$ comme le moment d'ordre n de la distribution de probabilité. Si la distribution de probabilité prend des valeurs non nulles seulement sur un intervalle fini, alors la connaissance de tous ses moments suffit pour retrouver toute la distribution.

Exercice 1.2.3 Démontrez que $\langle (u - \langle u \rangle)^2 \rangle = \langle u^2 \rangle - \langle u \rangle^2$ et ce quelle que soit la probabilité P . Ce résultat est très important en pratique et il est utilisé souvent.

Exercice 1.2.4 Considérons un jeu de Loto où la probabilité de perdre 1\$ (le coût du billet) est de 999/1000 et la probabilité de gagner 500\$ est de 1/1000. Calculez le gain moyen et l'écart type correspondant. Est-ce que les gens participent aux jeux de hasard lorsque les fluctuations sont grandes ou petites et pourquoi?

1.3 Analyse combinatoire

Additionner les probabilités d'événements complémentaires va parfois nous forcer à compter vite! Alors cette section va nous donner quelques outils pour compter vite.

1.3.1 Permutations, arrangements et combinaisons

Si j'ai N objets distincts, combien y a-t-il de façons différentes de les mettre en ordre? Par exemple, prenons trois objets, A,B,C. On peut les mettre en ordre des six façons suivantes:

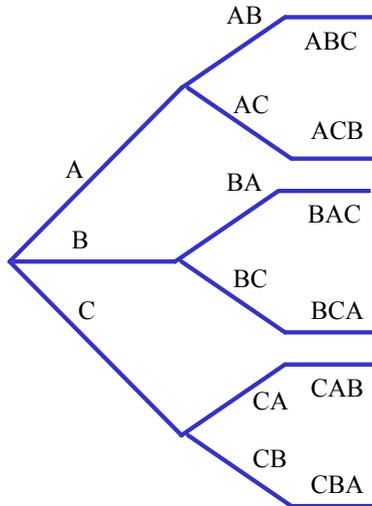


Figure 1-2 Arbre des possibilités illustrant pour $N = 3$ la façon dont on construit toutes les permutations possibles de N objets. À la première étape il y a ici $N = 3$ choix possibles, puis $N(N - 1) = 6$ à la deuxième étape et finalement $N(N - 1)(N - 2) = 6$ à la dernière étape.

ABC, ACB, BAC, BCA, CAB, CBA

En général, il y a N façons de choisir le premier objet, $N - 1$ façons de choisir le deuxième etc. Le résultat total est noté $N!$:

$$N! \equiv N \times (N - 1) \times (N - 2) \times \dots \times 2 \times 1 \quad (1.7)$$

Définition 5 Chacun des ordres différents de N objets est appelé une “permutation”. On dit donc qu’il y a “ N factorielle” permutations de N objets différents.

Pour se convaincre que la formule ci-dessus pour le nombre de permutations est la bonne, il suffit de traiter par exemple le cas de trois objets ABC ci-dessus à l’aide d’un schéma, comme à la Fig.(1-2), illustrant tous les choix possibles.

Définition 6 On veut savoir le nombre de façons d’ordonner n objets tirés de N . Si chaque ordre des n objets pris parmi N est considéré comme différent, on parle alors du nombre “d’arrangements” de N objets pris n à la fois.

En utilisant le même raisonnement que pour les permutations, en particulier l’arbre des possibilités, on voit que le nombre d’arrangements de n objets pris parmi N est égal à

$$N(N - 1) \dots (N - n + 1) = \frac{N!}{(N - n)!} \quad (1.8)$$

Exemple 7 On veut savoir combien de conseils étudiants différents il est possible de former dans une classe de 30 étudiants. Le conseil étudiant est formé d’un président, d’un vice-président, d’un secrétaire et d’un trésorier. La réponse est qu’il y a $30!/26! = 30 * 29 * 28 * 27$ possibilités.

Maintenant un peu plus compliqué. Parmi N objets différents, je veux savoir combien il est possible de faire de sous-groupes différents composés de n objets lorsque je ne me préoccupe pas de l’ordre dans lequel les n objets ont été choisis.

La réponse est qu'il suffit de prendre le résultat précédent pour les arrangements et de diviser par $n!$ puisque chaque permutation de n objets était comptée comme différente dans les *arrangements*. En d'autres mots, on obtient

$$C_n^N \equiv \binom{N}{n} \equiv \frac{N!}{n!(N-n)!} \quad (1.9)$$

Ce rapport revient si souvent qu'on prend la peine de lui assigner un symbole, C_n^N ou encore $\binom{N}{n}$.

Exemple 8 Dans notre exemple de conseil étudiant, si on ne se préoccupe pas de qui occupe chacun des postes, il faut diviser le nombre d'arrangements par $4!$ puisque c'est le nombre de façons dont les quatre étudiants du conseil peuvent occuper chacun des postes.

Définition 9 Le nombre de façons de choisir n objets parmi N lorsqu'on ne se préoccupe pas de l'ordre dans lequel les n objets ont été choisis est appelé le nombre de "combinaisons" de n objets parmi N .

Remarque 6 Voici une autre façon d'arriver au résultat pour le nombre de combinaisons. Mettons d'abord les N objets en ordre. Il y a $N!$ façons de faire cela. Divisons chacun des $N!$ groupes de N objets en deux sous-groupes, un contenant les n premiers objets et l'autre les $N - n$ suivants. Dans les $N!$ séquences que j'ai formé au départ, il arrivera plusieurs fois que les n premiers objets seront les mêmes sauf pour l'ordre dans lequel ils apparaîtront. Il y a $n!$ façons de les ordonner. Pour chacun de ces ordres, il y a $(N - n)!$ façons d'ordonner les objets restant. Chacun des différents groupes de n objets, est donc répété $n!(N - n)!$ fois lorsqu'on compte toutes les permutations possibles des deux sous-groupes. Le nombre de groupes différents de n objets choisis parmi N sera donc $\frac{N!}{n!(N-n)!}$.

Exemple 10 On vérifie aisément qu'il y a trois façons de faire des groupes de deux objets lorsque ceux-ci sont choisis parmi trois. Ceci se vérifie explicitement à l'aide de l'exemple A, B, C donné plus haut ainsi qu'à l'aide de $3! / (2!1!) = 3 \times 2 / 2 = 3$.

Exercice 1.3.1 Prenez quatre objets différents, disons A, B, C, D . Écrivez au long les 24 façons de les ordonner et identifiez le nombre de groupes de deux objets qu'on peut former. Montrez que l'on retrouve bien le résultat que l'on déduirait de C_2^4 .

Exercice 1.3.2 Nous sommes dans une réunion de vingt personnes. Chacune donne la main à toutes les autres à l'arrivée. Combien de poignées de mains différentes seront échangées? Résolvez ce problème de deux façons différentes.

Exercice 1.3.3 À partir de N objets, je désire faire deux groupes différents. Le premier groupe contiendra n_1 objets, et le deuxième n_2 objets. Combien de possibilités y a-t-il? Montrez que si je veux faire i groupes contenant respectivement n_1, n_2, \dots jusqu'à n_i objets, alors dans le cas général il y a

$$\frac{N!}{n_1!n_2!\dots n_i!(N - n_1 - n_2 - \dots - n_i)!} \quad (1.10)$$

façons de faire cela.

1.3.2 Binôme de Newton

Dans cette section, nous démontrerons un petit théorème qui nous sera très utile dans la section suivante. En fait, même la dernière technique de preuve sera en partie réutilisée plus tard!

Les nombres C_n^N sont aussi connus sous le nom de “coefficients du binôme”. Ce nom vient d’une formule découverte par Newton, ”le *binôme de Newton*”. Cette formule, que nous trouverons très utile plus tard, permet de développer rapidement un binôme $p + q$ à une puissance quelconque:

$$(p + q)^N = \sum_{n=0}^N \frac{N!}{n!(N-n)!} p^n q^{N-n} \quad (1.11)$$

Par définition

$$0! = 1$$

On peut prouver le résultat général binôme de Newton de plusieurs façons.

Preuve 1: Considérez chacun des N termes $(p + q)$ comme s’ils étaient nos N objets. Il faut trouver combien il y a de façons d’obtenir le monôme $p^n q^{N-n}$ lorsqu’on fait le produit des N objets. C’est la même chose que de chercher le nombre de façons dont on peut diviser le groupe de N objets en deux sous-groupes, celui des p et celui des q , contenant chacun respectivement n et $N - n$ objets. Il y a C_n^N façons différentes de faire ces deux sous-groupes. Une façon est considérée différente si au moins un des N termes $(p + q)$ n’appartient pas au même groupe (p ou q).

Preuve 2: On peut prouver la formule du binôme par induction. Elle est facile à vérifier pour $N = 1$ et $N = 2$. Si la formule est vraie pour N , alors on peut vérifier qu’elle est vraie pour $N + 1$. (Exercice!)

Preuve 3: Nous allons généraliser le truc suivant. Supposons qu’on ait un polynôme,

$$f(x) = a + bx + cx^2 + dx^3. \quad (1.12)$$

Je peux trouver la valeur du coefficient de x^2 en dérivant deux fois le polynôme et en évaluant le résultat à $x = 0$. Soyons plus explicite. En dérivant deux fois je trouve

$$\frac{1}{2} \frac{d^2 f(x)}{dx^2} = c + 3dx \quad (1.13)$$

puis en évaluant à $x = 0$, j’obtiens

$$\left. \frac{1}{2} \frac{d^2 f(x)}{dx^2} \right|_{x=0} = c \quad (1.14)$$

On peut généraliser cette approche de la façon suivante. On veut trouver les coefficients a_n^N , supposés inconnus, du polynôme suivant

$$(p + q)^N = \sum_{k=0}^N a_k^N p^k q^{N-k} \quad (1.15)$$

On note l’identité suivante

$$\left. \frac{1}{n!} \frac{\partial^n}{\partial p^n} p^k \right|_{p=0} = \begin{cases} 1 & \text{si } k = n \\ 0 & \text{si } k < n \end{cases} \quad (1.16)$$

La dérivée courbée indique une dérivée partielle, un concept que nous rencontrerons souvent dans ce cours. Lorsqu’on prend la dérivée partielle par rapport à p on dérive par rapport à p en gardant toutes les autres variables constantes, en particulier ici q . Pour trouver a_n^N il suffit d’appliquer l’opérateur $\frac{1}{n!} \frac{\partial^n}{\partial p^n} \frac{1}{(N-n)!} \frac{\partial^{N-n}}{\partial q^{N-n}}$

à $(p+q)^N$ puisque appliqué au membre de droite de l'Éq.(1.15) cet opérateur donnera la valeur zéro lorsqu'il agit sur un monôme $p^k q^{N-k}$ où $k \neq n$, alors qu'il donnera a_n^N lorsque $k = n$, comme on peut le voir à partir de l'Éq.(1.16). On trouve donc le résultat recherché,

$$a_n^N = \left(\frac{1}{n!} \frac{\partial^n}{\partial p^n} \right) \left(\frac{1}{(N-n)!} \frac{\partial^{N-n}}{\partial q^{N-n}} \right) (p+q)^N \quad (1.17)$$

$$= \frac{N!}{n!(N-n)!} = C_n^N \quad (1.18)$$

1.4 Exemples de distributions de probabilité

1.4.1 La marche aléatoire: distribution de probabilité binomiale⁶

Pour simplifier la discussion, on commence par la marche aléatoire en une dimension. Afin d'éviter les exemples d'alcoolique, considérons plutôt une particule qui se déplace en une dimension d'une distance ℓ puis fait une collision. Suite à cette collision, la particule repart à droite avec une probabilité p ou à gauche avec une probabilité q . On voit qu'il s'agit là d'un modèle d'application très générale en physique. Que ce soit pour la propagation des particules dans un gaz, ou de la lumière dans un milieu aléatoire, ou du mouvement Brownien, de simples généralisations du problème de base nous permettront de traiter de plusieurs cas. Dans le langage des probabilités, p et q sont les probabilités à priori. En particulier,

$$p + q = 1 \quad (1.19)$$

puisque aller à gauche ou à droite donne toutes les possibilités après une collision. On veut calculer la probabilité de se retrouver à la distance

$$x = m\ell \quad (1.20)$$

après N pas.

En suivant nos règles élémentaires pour l'addition et la multiplication des probabilités, on peut facilement déduire le résultat suivant: La probabilité de faire un pas à droite, puis un pas à gauche est donnée par pq (règle du *ET*). Celle de faire un à gauche puis un à droite par qp . Celle d'avoir fait un pas à gauche et un à droite dans n'importe quel ordre (c'est à-dire avec à droite en premier *OU* à gauche en premier) est donné par la somme des probabilités pour les deux cas précédents (règle du *OU*) soit $2pq$.

On généralise ce raisonnement à un cas arbitraire. La probabilité de faire une marche aléatoire particulière avec disons n_1 pas à droite puis $N - n_1$ pas à gauche est $p^{n_1} q^{N-n_1}$ (c'est la règle du *ET*). Si on ne se préoccupe pas de l'ordre dans lequel les pas à droite et à gauche ont été pris mais qu'on ne veut que connaître la probabilité $W_N(n_1)$ qu'au total n_1 pas à droite et $N - n_1$ pas à gauche aient été faits, il nous suffit d'additionner (règle du *OU*) toutes les façons de faire n_1 pas à droite lorsqu'on fait N pas en tout soit,

$$W_N(n_1) = \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} \quad (1.21)$$

Voilà! Cette expression, ainsi que la relation entre p et q Éq.(1.19) s'appelle la *distribution de probabilité binomiale*.

⁶Reif, Secs.1.1 et 1.2

La probabilité d'être rendu à une distance x se calcule aisément par un changement de variable.

$$x = (n_1 - (N - n_1))\ell \quad (1.22)$$

$$n_1 = \frac{1}{2} \left(N + \frac{x}{\ell} \right) \quad (1.23)$$

Substituant dans notre expression pour la probabilité d'obtenir n_1 nous avons

$$P_N(x) = W_N(n_1) = \frac{N!}{\frac{1}{2}(N + \frac{x}{\ell})! \frac{1}{2}(N - \frac{x}{\ell})!} p^{\frac{1}{2}(N + \frac{x}{\ell})} (1-p)^{\frac{1}{2}(N - \frac{x}{\ell})} \quad (1.24)$$

En d'autres mots, les changements de variable sont une affaire triviale pour les variables discrètes. Pour les variables continues, il faut faire plus attention comme nous le verrons plus loin.

Calcul des valeurs moyennes et de l'écart type pour la binomiale⁷

Nous devrions déjà commencer à développer un réflexe. On me donne une distribution de probabilité. Voilà, je vérifie qu'elle est normalisée et si ce n'est pas le cas je retourne l'objet pour défaut de manufacture. Puis, avec une distribution normalisée, je calcule quelques valeurs moyennes et l'écart-type, question de me donner une impression plus générale de cet objet. Allons-y.

Normalisation: C'est un petit calcul facile. J'utilise le binôme de Newton Éq.(1.11)

$$\sum_{n_1=0}^N W_N(n_1) = \sum_{n_1=0}^N \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} = (p+q)^N = 1 \quad (1.25)$$

La dernière égalité vient de la condition de normalisation $p+q=1$.

Valeur moyenne: Il faut calculer

$$\bar{n}_1 = \sum_{n_1=0}^N n_1 W_N(n_1) = \sum_{n_1=0}^N n_1 \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} \quad (1.26)$$

Pas facile à faire apparemment! Il y a un petit truc qui va nous sortir du pétrin. On utilise continuellement ce petit truc pour trouver une nouvelle intégrale lorsqu'une est connue en fonction d'un paramètre. Nous le ferons pour la gaussienne. Revenant aux sommes, on constate que la somme donnant la condition de normalisation est connue en fonction des paramètres p et q quels qu'ils soient, c'est-à-dire même s'ils ne satisfont $p+q=1$. Nous allons utiliser ceci à notre avantage. En effet,

$$n_1 p^{n_1} = p \frac{\partial}{\partial p} (p^{n_1}) \quad (1.27)$$

où la dérivée courbée reviendra souvent. Cette dérivée courbée est la dérivée partielle que nous avons déjà rencontrée à la section précédente. Comme toujours, elle veut dire qu'on dérive par rapport à la variable indiquée, ici p , en gardant toutes les autres variables constantes. Ainsi, on ne dérive pas q par rapport à p . Celles-ci sont *pour le moment* considérées indépendantes. Utilisant ce simple résultat, on obtient donc,

$$\bar{n}_1 = \sum_{n_1=0}^N \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} n_1 = \sum_{n_1=0}^N \frac{N!}{n_1!(N-n_1)!} p \frac{\partial}{\partial p} (p^{n_1}) q^{N-n_1} \quad (1.28)$$

⁷Reif, Sec.1.4

$$= p \frac{\partial}{\partial p} \left[\sum_{n_1=0}^N \frac{N!}{n_1! (N-n_1)!} (p^{n_1}) q^{N-n_1} \right] = p \frac{\partial}{\partial p} (p+q)^N = pN (p+q)^{N-1} \quad (1.29)$$

Comme ce résultat est valable pour p et q arbitraires, substituons $q = 1 - p$ pour obtenir,

$$\boxed{\bar{n}_1 = Np} \quad (1.30)$$

Remarque 7 Le truc que nous venons d'utiliser, c'est-à-dire dériver la condition de normalisation $(p+q)^N$ par rapport à p pour trouver la valeur moyenne, est un truc d'une très grande généralité. Non seulement peut-on l'utiliser comme ci-dessous pour calculer l'écart type, on utilisera un truc tout à fait analogue pour calculer moyenne et écart type de la gaussienne et (en devoir) de la distribution de Poisson. En physique statistique, la fonction de partition Z jouera le rôle de $(p+q)^N$ et les dérivées de la fonction de partition (de son logarithme) nous donneront des valeurs moyennes de quantités observables, comme l'énergie disons. Nous venons donc d'apprendre un truc d'une très très grande utilité. C'est un bon investissement de faire tous les efforts ici pour le comprendre!

Écart type: C'est pas juste, vous connaissez tous les trucs! Après un peu d'amusement mathématique, on obtient

$$\boxed{(\Delta n_1)^2 \equiv n_1^2 - \bar{n}_1^2 = Npq} \quad (1.31)$$

Pour imaginer tout ça, voici à la Fig.(1-3) un exemple explicite de distribution de probabilité binomiale. La valeur moyenne ici est $Np = 4$ et l'écart type $\sqrt{Npq} = \sqrt{2.4} = 1.55$.

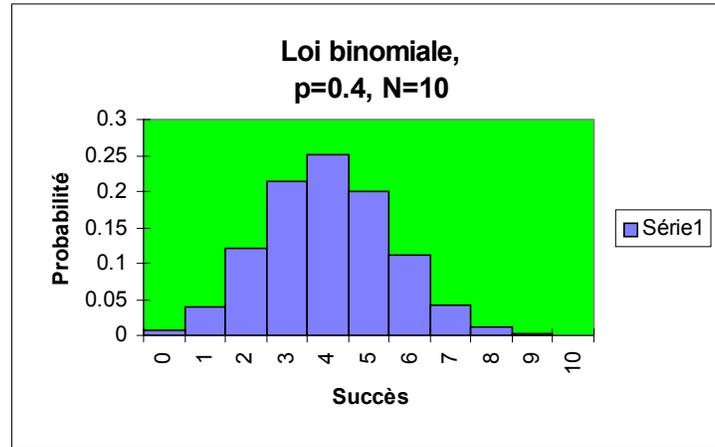


Figure 1-3 Distribution binomiale de valeur moyenne 4 et d'écart type 2.4.

Un peu de résultats qualitatifs: Une chose importante à remarquer est que la largeur relative de la distribution de probabilité, c'est-à-dire

$$\frac{\sqrt{n_1^2 - \bar{n}_1^2}}{\bar{n}_1} = \frac{\sqrt{Npq}}{Np} = \sqrt{\frac{q}{p}} \frac{1}{\sqrt{N}} \quad (1.32)$$

diminue avec le nombre de pas pris au total, N . C'est un résultat général des statistiques. Lorsqu'une variable aléatoire est la somme d'un grand nombre d'autres variables aléatoires indépendantes, l'incertitude relative sur la variable somme diminue avec la racine carrée du nombre de variables aléatoires dont elle est la somme. Ici, on peut considérer la distance parcourue comme un somme de N variables aléatoires. Nous reviendrons à cette discussion en fin de chapitre.

1.4.2 Distribution de probabilité gaussienne

La petite discussion innocente de la fin de la section précédente a en fait de grandes répercussions. En effet, ce rétrécissement relatif de la distribution de probabilité implique que pour $N \rightarrow \infty$ nous obtiendrons éventuellement toujours la même distribution de probabilité. Ce n'est pas encore évident, mais regardons d'abord en image ce qui se passe. Nous allons faire des graphiques de la distribution de probabilité binomiale pour différentes valeurs de N . Afin de nous concentrer toujours autour de la région où la probabilité est maximale, nous allons mettre notre origine à Np . De plus, nous savons que la largeur de la distribution de probabilité augmentera comme \sqrt{Npq} . Nous contracterons donc l'abscisse par ce facteur de telle sorte que nous ayons toujours la même vue relative. En d'autres mots, nous tracerons en abscisse $(n - Np) / \sqrt{Npq}$. La valeur de la probabilité au maximum décroît aussi avec N car dans un intervalle \sqrt{Npq} on retrouve la majeure partie de la probabilité totale. Comme le maximum n'est qu'un des points, sa probabilité décroîtra proportionnellement à \sqrt{Npq} . Nous porterons donc en ordonnée $P(n) \sqrt{Npq}$ afin de garder la hauteur de cet axe à peu près fixe. Nous ne faisons pas ces changements d'échelle pour $N = 1$ afin d'illustrer la probabilité de départ tel quel. Les quatre figures suivantes, Figs.(1-4) à (1-7), illustrent donc la distribution de probabilité binomiale avec les changements d'échelle que nous venons de discuter en abscisse et en ordonnée.

Nous pouvons voir que nous atteignons une distribution limite. En d'autres mots, la forme de la distribution de probabilité, lorsque tracée avec les axes que nous nous sommes donnés, ne dépend plus de N . C'est le théorème de la limite centrale. L'outil moderne pour traiter analytiquement ce genre de genre théorème s'appelle le groupe de renormalisation.⁸ Cette méthode se retrouve dans la théorie moderne des transition de phase. Plutôt que de se lancer dans les grands théorèmes tout de suite, voyons analytiquement ce qui se passe effectivement avec notre binomiale. On commence la discussion par deux petits préliminaires mathématiques.

Développement asymptotique pour $N!$ et approximation pour N grand⁹

Considérons la limite N grand. On peut alors utiliser l'approximation suivante, connue sous le nom de formule de Stirling, que nous prouvons à la section 1.7:

$$\lim_{N \text{ grand}} N! \approx \exp\left(N \ln N - N + \frac{1}{2} \ln(2\pi N)\right) \quad (1.33)$$

$$\approx \sqrt{2\pi N} \exp(N \ln N - N) \quad (1.34)$$

Remarque 8 *En attendant la preuve formelle de ce résultat, présentée à fin du chapitre, on peut se convaincre de sa validité approximative à partir de l'argument suivant.*

$$\ln N! = \sum_{n=1}^N \ln n \simeq \int_1^N (\ln x) dx \simeq N \ln N - N \quad (1.35)$$

Remarque 9 *La formule de Stirling marche très bien même pour des N relativement petits, comme on peut s'en convaincre à partir des résultats numériques suivants: Même pour $N = 1$ on a .92214 à partir de la formule approximative. Pour $N > 10$ on a déjà environ 1% de précision, comme le montre le tableau suivant.*

⁸La distribution limite s'appelle un point fixe dans le jargon du groupe de renormalisation.

⁹Reif, Sec. 1.5

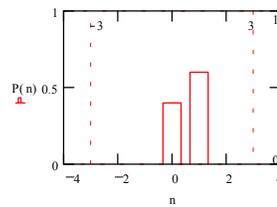
Binomiale pour différentes valeurs de N (Convergence vers la gaussienne)

Créé avec Mathcad 4.0. Fichier: bin-gaus.mcd

$$p := 0.6 \quad q := 1 - p$$

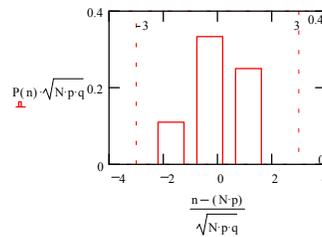
$$N := 1 \quad n := 0, 1.. N$$

$$P(n) := \frac{N!}{n! (N - n)!} p^n q^{N-n}$$



$$N := 2 \quad n := 0, 1.. 2$$

$$P(n) := \frac{N!}{n! (N - n)!} p^n q^{N-n}$$



$$N := 3 \quad n := 0, 1.. 3$$

$$P(n) := \frac{N!}{n! (N - n)!} p^n q^{N-n}$$

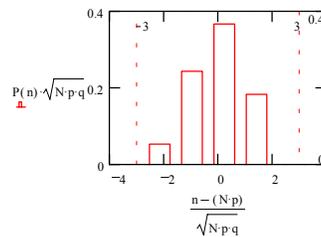
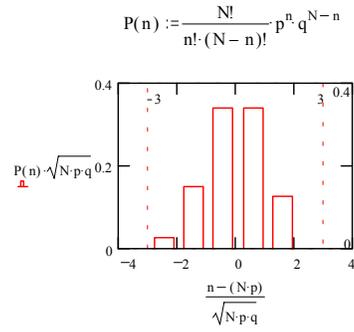
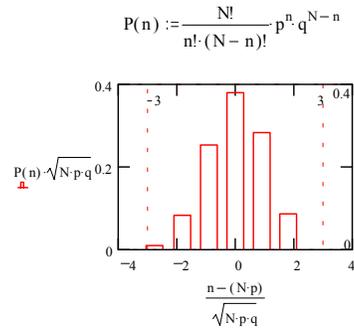


Figure 1-4 Convergence de la binomiale vers la gaussienne. $N = 1, 2, 3$

$N := 4 \quad n := 0, 1.. 4$



$N := 5 \quad n := 0, 1.. 5$



$N := 6 \quad n := 0, 1.. 6$

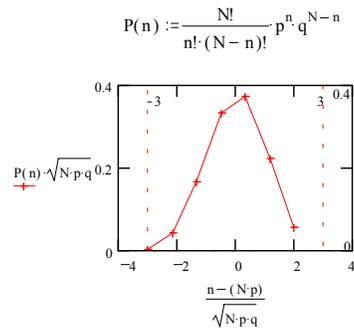
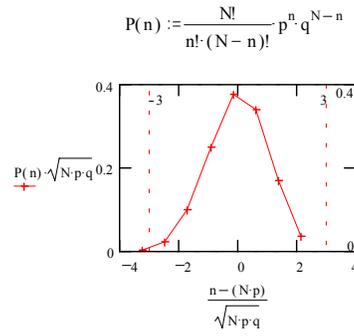
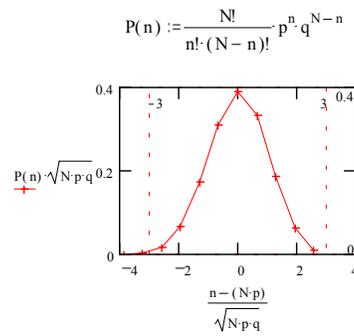


Figure 1-5 Convergence de la binomiale vers la gaussienne. $N = 4, 5, 6$

$N := 7 \quad n := 0, 1, \dots, 7$



$N := 10 \quad n := 0, 1, \dots, 10$



$N := 20 \quad n := 5, 6, \dots, 20$

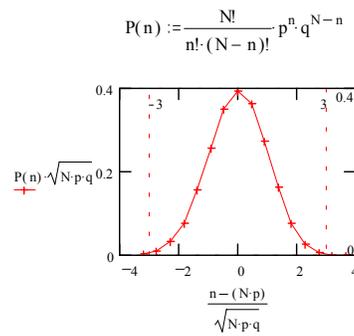
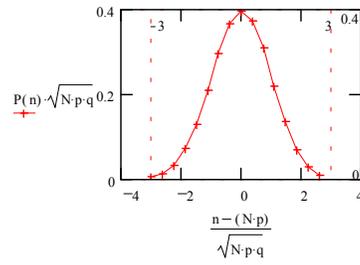


Figure 1-6 Convergence de la binomiale vers la gaussienne. $N = 7, 10, 20$

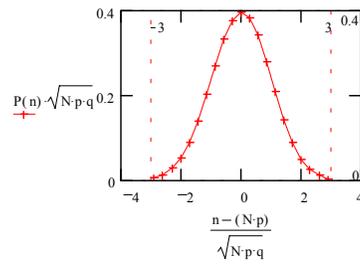
$N := 30 \quad n := 10, 11, 25$

$$P(n) := \frac{N!}{n!(N-n)!} p^n q^{N-n}$$



$N := 50 \quad n := 20, 21, 40$

$$P(n) := \frac{N!}{n!(N-n)!} p^n q^{N-n}$$



$N := 100 \quad n := 45, 46, 75$

$$P(n) := \frac{N!}{n!(N-n)!} p^n q^{N-n}$$

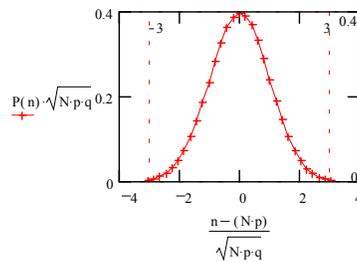


Figure 1-7 Convergence de la binomiale vers la gaussienne. $N = 30, 50, 100$

N	$N!$	$\exp(N \ln N - N + \frac{1}{2} \ln(2\pi N))$
2	2	1.919
3	6	5.8362
4	24	23.506
5	120	118.02
6	720	710.08
7	5040	4980.4
8	40,320	39,902
9	3.6288×10^5	3.5954×10^5
10	3.6288×10^6	3.5987×10^6

Quelle est la meilleure stratégie pour trouver une approximation analytique pour la binomiale lorsque N est grand? C'est évidemment d'organiser un développement de Taylor qui utilise la petitesse de $1/N$. Rappelons-nous que c'est autour du maximum de $W_N(n_1)$ qu'il faut travailler et que ce maximum est situé à $n_1 = Np$, c'est à dire à une valeur de n_1 qui est d'ordre N . Une approche directe serait de faire un développement de Taylor autour du maximum. Nous verrons que ce n'est pas vraiment une bonne stratégie.

Utilisant

$$p^n = e^{n \ln p} \quad (1.36)$$

ainsi que la formule de Stirling pour expliciter un peu la forme de $W_N(n_1)$ on voit que

$$W_N(n_1) = \exp(f_N(n_1)) \quad (1.37)$$

où $f(n_1)$ contient des termes d'ordre n_1 , ou $n_1 \ln n_1$ ou $\ln n_1$. Pour être plus explicite, écrivons au long,

$$\begin{aligned} W_N(n_1) &= \left[\frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} \right] \approx \\ &\exp \left[\ln(p^{n_1} q^{N-n_1}) + N \ln N - \underline{N} + \frac{1}{2} \ln(2\pi N) - n_1 \ln n_1 + \underline{n_1} - \frac{1}{2} \ln(2\pi n_1) \right. \\ &\quad \left. - (N-n_1) \ln(N-n_1) + \underline{(N-n_1)} - \frac{1}{2} \ln(2\pi(N-n_1)) \right] \end{aligned} \quad (1.38)$$

Les termes soulignés se simplifient. Regroupons ensuite les termes d'ordre le plus bas en n_1 , soit $\ln(n_1)$. On a alors, utilisant $\exp(\frac{1}{2} \ln x) = \sqrt{x}$,

$$W_N(n_1) = \sqrt{\frac{2\pi N}{2\pi n_1 2\pi(N-n_1)}} \exp[n_1 \ln p + (N-n_1) \ln q] \quad (1.39)$$

$$+ N \ln N - n_1 \ln n_1 - (N-n_1) \ln(N-n_1)] \quad (1.40)$$

$$\equiv \sqrt{\frac{N}{2\pi n_1(N-n_1)}} \exp g_N(n_1) \quad (1.41)$$

La dernière équation définit $g_N(n_1)$. Le préfacteur est une fonction algébrique de n_1 dont la valeur autour de $n_1 = Np$ ne varie que sur des intervalles d'ordre N . Le facteur exponentiel est évidemment beaucoup plus sensible aux variations de n_1 . Commençons donc par développer l'argument de l'exponentielle $g_N(n_1)$ autour du maximum $\partial g(n_1)/\partial n_1|_{n_1^{\max}} = 0$. Le maximum de l'exponentielle est situé au même endroit que le maximum de son argument. Le développement prend la forme

$$g_N(n_1) = g_N(n_1^{\max}) + \frac{1}{2} \frac{d^2 g_N(n_1)}{dn_1^2} (n_1 - n_1^{\max})^2 + \dots \quad (1.42)$$

la valeur de n_1^{\max} étant donnée par la solution de

$$\left. \frac{\partial g_N(n_1)}{\partial n_1} \right|_{n_1^{\max}} = \ln p - \ln q - \ln n_1^{\max} + \ln(N - n_1^{\max}) = 0 \quad (1.43)$$

et la dérivée seconde étant obtenue de

$$\left. \frac{\partial^2 g_N(n_1)}{\partial n_1^2} \right|_{n_1^{\max}} = \left[-\frac{1}{N - n_1} - \frac{1}{n_1} \right] \Big|_{n_1^{\max}} \quad (1.44)$$

On remarque que chaque dérivée successive est plus petite que la précédente par un facteur $1/n_1^{\max} \propto 1/N$. La série de Taylor Éq.(1.42) converge donc rapidement.

Il ne reste plus qu'à évaluer chacun des termes entrant dans la série de Taylor Éq.(1.42) pour l'argument de l'exponentielle. La solution de l'Éq.(1.43) pour le maximum s'obtient à partir de

$$\ln\left(\frac{p}{q}\right) = \ln\left(\frac{n_1^{\max}}{N - n_1^{\max}}\right) \quad (1.45)$$

dont la solution, utilisant $q = 1 - p$, est $n_1^{\max} = Np$. Substituant ce résultat dans l'expression pour la dérivée seconde Éq.(1.44), on a

$$\left[-\frac{1}{N - n_1} - \frac{1}{n_1} \right] \Big|_{n_1^{\max}} = -\frac{1}{N} \left(\frac{1}{1 - p} + \frac{1}{p} \right) = -\frac{1}{N} \frac{1}{pq} \quad (1.46)$$

alors que le terme

$$g_N(n_1) = [n_1 \ln p + (N - n_1) \ln q] \quad (1.47)$$

$$+ N \ln N - n_1 \ln n_1 - (N - n_1) \ln(N - n_1) \quad (1.48)$$

s'annule pour $n_1 = n_1^{\max}$ comme on peut le voir à partir de

$$-n_1^{\max} \ln n_1^{\max} - (N - n_1^{\max}) \ln(N - n_1^{\max}) \quad (1.49)$$

$$= -Np \ln Np - Nq \ln Nq = -N \ln N - Np \ln p - Nq \ln q \quad (1.50)$$

Approximation pour $W_N(n_1)$: Substituant le résultat de tous ces calculs pour l'argument de l'exponentielle Éq.(1.42) dans l'équation de départ Éq.(1.41) on obtient

$$\lim_{N \text{ grand}} W_N(n_1) \approx \sqrt{\frac{N}{2\pi n_1(N - n_1)}} \exp\left(-\frac{1}{2} \frac{(n_1 - Np)^2}{Npq}\right). \quad (1.51)$$

Lorsque $n_1 = Np \pm \sqrt{Npq}$, l'exponentielle chute d'un facteur $e^{-1/2}$ par rapport à sa valeur au maximum. Par contre, la valeur du préfacteur n'a changé que par un facteur d'ordre $(1 + \mathcal{O}(1/\sqrt{N}))$ ce qui est négligeable. C'est donc une excellente approximation d'évaluer le préfacteur au maximum, ce qui nous laisse la forme asymptotique suivante à l'ordre dominant

$$\lim_{N \text{ grand}} W_N(n_1) \approx \frac{1}{\sqrt{2\pi Npq}} \exp\left(-\frac{1}{2} \frac{(n_1 - Np)^2}{Npq}\right) \quad (1.52)$$

$$\lim_{N \text{ grand}} W_N(n_1) \approx \frac{1}{\sqrt{2\pi(\Delta n_1)^2}} \exp\left(-\frac{1}{2} \frac{(n_1 - \bar{n}_1)^2}{(\Delta n_1)^2}\right) \quad (1.53)$$

Remarque 10 *La dernière façon d'écrire la distribution de probabilité dans cette limite suggère que le résultat est plus général que sa dérivation. Nous reviendrons là-dessus.*

Remarque 11 *Comment justifie-t-on d'avoir négligé les autres termes dans le développement de l'argument de l'exponentielle? En prenant une dérivée additionnelle de l'Eq.(1.44) on voit que le terme suivant dans l'exponentielle aurait eu la forme

$$\sim \frac{1}{N^2} (n_1 - Np)^3 \quad (1.54)$$

or, pour $n_1 = Np + \sqrt{Npq}$, déjà $W_N(n_1)$ est réduit d'un facteur $\exp(-1/2)$ par rapport à sa valeur au maximum, ce qui n'est qu'une autre manifestation de l'étroitesse relative de la distribution de probabilité. Le terme suivant n'ajouterait qu'une correction d'ordre

$$\frac{1}{N^2} (\sqrt{Npq})^3 \sim \frac{1}{\sqrt{N}} \quad (1.55)$$

à l'argument de l'exponentielle, i.e. on trouverait que $W_N(n_1)$ est réduit d'un facteur $\exp\left(-1/2 + O\left(1/\sqrt{N}\right)\right)$. On néglige donc ces corrections dans la limite $N \rightarrow \infty$. Pour $n_1 = Np + O(N)$ la binomiale originale est très différente de la gaussienne, mais par contre la valeur de la probabilité est tellement petite ($\sim \exp(-N)$) que ces corrections ne peuvent changer aucune valeur moyenne de façon importante.

Remarque 12 *Cela aurait été une très mauvaise idée de développer directement l'exponentielle plutôt que de développer son argument $g_N(n_1)$. En effet, soit

$$\exp\left(-\frac{a}{N} (n_1 - Np)^2 + \frac{b}{N^2} (n_1 - Np)^3 + \dots\right) \quad (1.56)$$

où a et b sont des constantes de l'ordre de l'unité. Même pour $n_1 = Np + 10\sqrt{N}$ le premier terme de l'argument de l'exponentielle $\frac{a}{N} (n_1 - Np)^2 \simeq 100a$ est une bonne approximation puisque le terme suivant dans l'argument est plus petit par un facteur $1/\sqrt{N}$, tel que mentionné dans la remarque précédente. Cependant, le développement de l'exponentielle elle-même en puissances de $(n_1 - Np)$ ne converge que pour une série très longue lorsque $\frac{a}{N} (n_1 - Np)^2 \simeq 100a$.

Remarque 13 *Symétrie : La distribution de probabilité gaussienne que nous avons obtenue est symétrique (paire) par rapport à la valeur moyenne et ce même si la distribution de départ, la binomiale, n'est pas symétrique autour de la moyenne lorsque $p \neq q$. On peut voir à partir de la remarque précédente que les premières contributions qui briseraient la symétrie par rapport à la valeur moyenne viennent du terme en $(n_1 - Np)^3$ et que ceux-ci sont d'ordre $1/N$ plus petits que le terme dominant. Ils sont donc négligeables dans la limite N grand. Dans le jargon du groupe de renormalisation, "l'asymétrie est non-pertinente". Cette "disparition" de l'asymétrie suggère que beaucoup de détails de la distribution de départ (ici la binomiale) n'ont pas d'importance pour la distribution obtenue dans la limite N grand. Le théorème de la limite centrale prouve que c'est le cas. Nous discuterons un peu plus en détails de ce théorème dans la section sur le marche aléatoire et la loi des grands nombres. Une preuve complète se trouve à la section 1.8.

Remarque 14 *Si l'écart type est suffisamment grand (ce qui est le cas si N est grand) cette forme approximative de la binomiale donne une distribution de probabilité qui est normalisée, à condition de pouvoir négliger des termes exponentiellement petits. En effet, en utilisant la formule de sommation de Poisson¹⁰

$$\sum_{n=-\infty}^{\infty} f(n) = \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} d\theta e^{2\pi i n \theta} f(\theta) \quad (1.57)$$

¹⁰Le Bellac, p.161

on peut prouver que pour $\overline{(\Delta n_1)^2} \gtrsim 1$

$$\sum_{n_1=-\infty}^{\infty} \exp\left(-\frac{1}{2} \frac{(n_1 - \overline{n_1})^2}{\overline{(\Delta n_1)^2}}\right) = \sqrt{2\pi \overline{(\Delta n_1)^2}} \left(1 + 2e^{-2\pi^2 \overline{(\Delta n_1)^2}} + \dots\right) \quad (1.58)$$

Limite du continuum et distribution de probabilité gaussienne¹¹

On sait déjà d'après les tracés de la binomiale que nous avons fait plus tôt que dans la limite $N \rightarrow \infty$ la probabilité ne change plus tellement lorsque n_1 change de n_1 à $n_1 + 1$. En effet, il faut que n_1 change beaucoup, soit de l'ordre de \sqrt{Npq} , pour que le changement soit appréciable. Les courbes suggèrent aussi que dans la limite $N \rightarrow \infty$ une version continue de la distribution de probabilité devient naturelle. La probabilité que n_1 soit compris entre n_1 et $n_1 + dn_1$ est alors simplement égale environ à $W_N(n_1) dn_1$. Par exemple, si $dn_1 = 4$ la probabilité est égale à quatre fois la probabilité évaluée à n_1 .

On peut rendre tout ça encore plus naturel lorsqu'il s'agit de décrire une vraie marche aléatoire. Pour un pas de grandeur ℓ à chaque déplacement, on a que le déplacement x est donné comme à l'équation Éq.(1.22) par

$$x = (n_1 - (N - n_1)) \ell \quad (1.59)$$

soit le nombre de pas à droite n_1 moins le nombre de pas à gauche $(N - n_1)$ fois le déplacement ℓ à chaque pas. Substituant alors

$$n_1 = \frac{1}{2} \left(N + \frac{x}{\ell}\right) \quad (1.60)$$

dans l'équation pour la binomiale Éq.(1.53) dans la limite $N \rightarrow \infty$ on obtient,

$$W_N(n_1) dn_1 = W_N(x) \frac{dx}{2\ell} \equiv \mathcal{P}(x) dx \quad (1.61)$$

$$\mathcal{P}(x) = \frac{1}{2\ell \sqrt{2\pi Npq}} \exp\left(-\frac{1}{2} \frac{\left(\frac{1}{2} \left(N + \frac{x}{\ell}\right) - Np\right)^2}{Npq}\right) \quad (1.62)$$

Multipliant numérateur et dénominateur de l'argument de l'exponentielle par $(2\ell)^2$, ce dernier résultat se réécrit avec les définitions

$$\boxed{\mu \equiv 2N\ell p - N\ell = N\ell(p - q)} \quad (1.63)$$

$$\boxed{\sigma \equiv 2\ell \sqrt{Npq}} \quad (1.64)$$

sous la forme générique suivante

$$\boxed{\mathcal{P}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}\right)} \quad (1.65)$$

Définition 11 La quantité $\mathcal{P}(x)$ est une densité de probabilité, c'est-à-dire qu'il faut multiplier par dx pour obtenir une probabilité.

Remarque 15 La forme particulière de $\mathcal{P}(x)$ trouvée ci-dessus s'appelle la distribution de probabilité gaussienne ou encore, en statistiques ou dans bien d'autres domaines, la distribution de probabilité normale.

¹¹Reif, Sec. 1.6

On s'attend à ce que μ soit la valeur moyenne du déplacement puisque c'est la valeur qu'on trouve en substituant $n_1 = Np$ dans l'équation Éq.(1.59). De même pour l'écart type, on s'attend à ce qu'il soit égal à σ puisque $dx = 2\ell dn_1$ et que l'écart type en n_1 est égal à \sqrt{Npq} . Puisque nous avons bien développé le réflexe de calculer valeur moyenne et écart type, vérifions ceci explicitement.

Moyennes calculées dans le continuum: Tout d'abord une remarque générale. Pour une probabilité quelconque $P(n_1)$, lorsqu'on passe dans le continuum, les sommes discrètes deviennent des intégrales, comme on peut le voir en utilisant la limite habituelle pour définir une somme à partir d'une intégrale

$$\sum_{n_1} P(n_1) \dots \rightarrow \int dx \mathcal{P}(x) \dots \quad (1.66)$$

Donc, en particulier pour n'importe quelle valeur moyenne d'une fonction de la variable aléatoire on a

$$\sum_{n_1} P(n_1) f(n_1) \rightarrow \int dx \mathcal{P}(x) f(x)$$

Normalisation: Pour vérifier la normalisation de la normale Éq.(1.62) nous devons vérifier que

$$\int dx \mathcal{P}(x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}\right) dx = 1. \quad (1.67)$$

Notez que les bornes ont été prises de $-\infty$ à ∞ parce que la distance la plus grande possible à droite est $x_{\max} = N\ell$ alors qu'elle est $-N\ell$ à gauche, ce qui donne bien les bornes d'intégration que nous avons choisies dans la limite $N \rightarrow \infty$.

Pour vérifier la normalisation, il suffit d'utiliser le truc suivant¹². On veut évaluer

$$\int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}} dx \quad (1.68)$$

Pour ce faire, on évalue le carré de cette intégrale et on prend la racine carré à la fin. En effet, en changeant en coordonnées polaires, le carré de l'intégrale est facile à faire, comme on peut le voir dans ce qui suit

$$\left[\int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}} dx \right]^2 = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}} dx \int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma^2}} dy \quad (1.69)$$

$$= \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}} e^{-\frac{y^2}{2\sigma^2}} dx dy \quad (1.70)$$

$$= \int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2\sigma^2}} dx dy \quad (1.71)$$

$$= \int_0^{\infty} r dr \int_0^{2\pi} d\theta e^{-\frac{r^2}{2\sigma^2}} \quad (1.72)$$

$$= 2\pi \int_0^{\infty} r dr e^{-\frac{r^2}{2\sigma^2}} \quad (1.73)$$

$$= 2\pi\sigma^2 \int_0^{\infty} \frac{r dr}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}} \quad (1.74)$$

$$= -2\pi\sigma^2 e^{-\frac{r^2}{2\sigma^2}} \Big|_0^{\infty} \quad (1.75)$$

$$= 2\pi\sigma^2 \quad (1.76)$$

¹²Reif, annexe A2

En prenant la racine carrée, on voit qu'on a prouvé que la condition de normalisation Éq.(1.67) est en effet satisfaite. (Il faut évidemment faire le changement de variable $x \rightarrow x + \mu$.)

Moyenne: Pour calculer la valeur moyenne, on procède par un simple changement de variable et on utilise ensuite la condition de normalisation et le fait que l'intégrale de $-\infty$ à ∞ d'une fonction impaire s'annule. En effet,

$$\int dx \mathcal{P}(x) x = \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}\right) dx \quad (1.77)$$

$$= \int_{-\infty}^{\infty} \frac{x+\mu}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(x)^2}{\sigma^2}\right) dx \quad (1.78)$$

$$= 0 + \mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(x)^2}{\sigma^2}\right) dx \quad (1.79)$$

$$= \mu$$

*Écart type:*¹³ On utilise le même truc qu'avec la binomiale, c'est-à-dire qu'on dérive le résultat qu'on a obtenu dans le calcul de la normalisation et on dérive par rapport à σ^2 . Plus précisément

$$\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \sqrt{2\pi\sigma^2} \quad (1.80)$$

$$\frac{\partial}{\partial \sigma} \int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} \frac{(x-\mu)^2}{\sigma^3} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{\partial}{\partial \sigma} (\sqrt{2\pi}\sigma) \quad (1.81)$$

ou encore

$$\int_{-\infty}^{\infty} (x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \sqrt{2\pi}\sigma^3$$

de telle sorte que

$$\int_{-\infty}^{\infty} (x-\mu)^2 \mathcal{P}(x) dx = \sigma^2 \quad (1.82)$$

tel que prévu, cette équation nous donne l'écart type σ en prenant la racine carré de la variance.

Voici à quoi ressemble une gaussienne, Fig.(1-8), ayant la même valeur moyenne et le même écart type que la binomiale illustrée à la figure (1-3).

*Probabilité cumulative et fonction erreur:*¹⁴ Pour se familiariser un peu avec la gaussienne, demandons-nous la question suivante. Soit une variable aléatoire x décrite par une distribution de probabilité normale de moyenne μ et d'écart type σ . Quelle est la probabilité que la variable x prenne des valeurs situées à $\pm\sigma$ de la valeur moyenne?

La réponse à cette question s'obtient en se souvenant que la probabilité que la variable prenne une valeur dans l'intervalle dx est donnée par $\mathcal{P}(x) dx$. Sur un intervalle qui n'est pas infinitésimal, il suffit donc d'intégrer. On trouve donc que la réponse à la question demandée s'obtient en évaluant l'intégrale suivante, ce qui

¹³Reif, Annexe A.4

¹⁴Reif, Annexe A.5

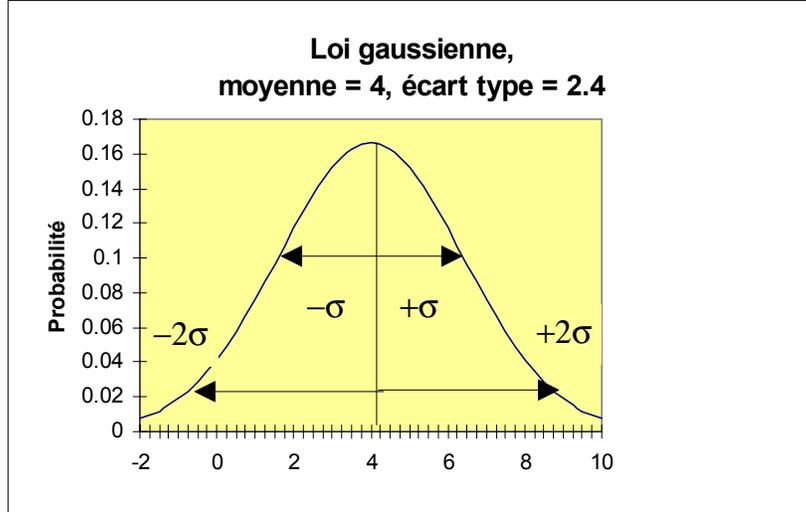


Figure 1-8 Distribution gaussienne de moyenne 4 et d'écart type 2.4.

se fait par des changements de variables simples

$$\int_{\mu-\sigma}^{\mu+\sigma} dx \mathcal{P}(x) = \int_{\mu-\sigma}^{\mu+\sigma} dx \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}\right) \quad (1.83)$$

$$= \int_{-\sigma}^{\sigma} dx \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(x)^2}{\sigma^2}\right) \quad (1.84)$$

$$= \int_{-1/\sqrt{2}}^{1/\sqrt{2}} dz \frac{1}{\sqrt{\pi}} \exp(-z^2) \quad (1.85)$$

$$= \frac{2}{\sqrt{\pi}} \int_0^{1/\sqrt{2}} dz \exp(-z^2) \quad (1.86)$$

$$= \operatorname{erf}\left(1/\sqrt{2}\right) = .68269 \quad (1.87)$$

où nous avons utilisé la définition de la fonction erreur, soit

$$\operatorname{erf}(y) = \frac{2}{\sqrt{\pi}} \int_0^y dz \exp(-z^2) \quad (1.88)$$

Cette fonction est tabulée et on connaît des développements asymptotiques dans différentes limites. La leçon importante est que, pour une gaussienne, environ les deux tiers des valeurs probables de x se trouvent à l'intérieur de $\pm\sigma$ de la moyenne. Mais attention, la chute est brutale. Pour $\pm 2\sigma$ on trouve $\operatorname{erf}(\sqrt{2}) = .9545$, et pour $\pm 3\sigma$, on a $\operatorname{erf}(3/\sqrt{2}) = .9973$, soit une probabilité plus faible que 3/1000 de se trouver à l'extérieur d'un intervalle $\pm 3\sigma$ de la moyenne.

Remarque 16 *Probabilité cumulative: En général on appelle la quantité*

$$\mathbb{P}(x) = \int_{-\infty}^x \mathcal{P}(x) dx \quad (1.89)$$

la "probabilité cumulative". Cette quantité $\mathbb{P}(x)$ est la probabilité (et non la densité de probabilité) pour que la variable aléatoire prenne une valeur plus petite que x . Dans le cas de la gaussienne, il est facile de voir que la probabilité cumulative est simplement reliée à la fonction erreur. Évidemment, $\mathbb{P}(x)$ est comprise entre 0 et 1.

1.5 Changements de variables pour les distributions de probabilité continues

Nous nous permettons une légère digression¹⁵ avant de retourner à notre problème principal, soit celui de la marche aléatoire. Nous nous posons le problème suivant. Soit une variable aléatoire continue u et une fonction de cette variable aléatoire $\phi(u)$. Sachant la densité de probabilité pour la variable u , quelle est celle pour la variable ϕ ? Pour répondre à cette question, il suffit de noter que la probabilité de tomber dans un intervalle du , soit $\mathcal{P}(u)du$ doit être égale à la probabilité de tomber dans l'intervalle correspondant $d\phi$. Plus spécifiquement,

$$\mathcal{W}(\phi) d\phi = \mathcal{P}(u) du \quad (1.90)$$

On serait donc porté à écrire

$$\mathcal{W}(\phi) = \mathcal{P}(u) \frac{du}{d\phi} ? \quad (1.91)$$

comme nous l'avons fait plus haut aux Éqs.(1.61)(1.59). Cependant, pour un changement de variable quelconque, il se peut que $du/d\phi$ soit négatif. Or, on veut que $\mathcal{W}(\phi)$ soit positif pour avoir l'interprétation d'une densité de probabilité. En choisissant toujours des $d\phi$ positifs pour l'intégration on trouve,

$$\boxed{\mathcal{W}(\phi) = \mathcal{P}(u) \left| \frac{du}{d\phi} \right|} \quad (1.92)$$

Pour être plus spécifique, prenons un exemple simple.

Exemple 12 Soit le changement de variable suivant avec $a > 0$,

$$\phi = -au \quad (1.93)$$

alors, en ramenant les bornes d'intégration sur $d\phi$ pour qu'elles soient du plus petit au plus grand et en utilisant la formule de changement de variable standard pour les intégrales, on retrouve le résultat recherché Éq.(1.92) comme le démontrent les manipulations qui suivent,

$$\int_{-1}^1 du \mathcal{P}(u) = \int_a^{-a} d\phi \frac{du}{d\phi} \mathcal{P}(u) \quad (1.94)$$

$$= \int_{-a}^a d\phi \left(-\frac{du}{d\phi} \right) \mathcal{P}(u) \quad (1.95)$$

$$= \int_{-a}^a d\phi \left| \frac{du}{d\phi} \right| \mathcal{P}(u) \quad (1.96)$$

$$= \int_{-a}^a d\phi \mathcal{W}(\phi) \quad (1.97)$$

Le changement de signe entre du et $d\phi$ est compensé par le changement de direction d'intégration. C'est pourquoi c'est toujours la valeur absolue de $\frac{du}{d\phi}$ qui intervient. Graphiquement, le tout peut être représenté comme à la figure (1-9). La probabilité est la même dans l'intervalle du ou dans l'intervalle $d\phi$ mais, pour $\phi = -2u$ la densité de probabilité $\mathcal{P}(u)$ pour u , est deux fois plus grande que celle pour ϕ , soit $\mathcal{P}(u) = 2 \times \mathcal{W}(\phi)$, parce que l'intervalle du correspondant à $d\phi$ est deux fois plus petit.

¹⁵Reif. Sec. 1.8

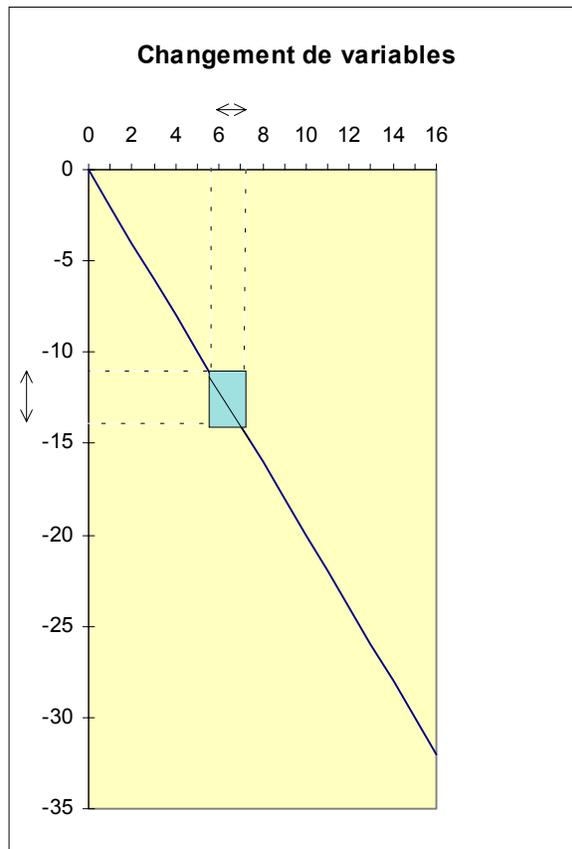


Figure 1-9 Changement de variables $\phi = -2u$, avec u sur l'axe horizontal, et ϕ sur l'axe vertical.

Et voilà. Il y a cependant une complication additionnelle encore. Il se peut que le changement de variable ne soit pas biunivoque. En d'autres mots, il se pourrait que deux valeurs différentes de u correspondent à la même valeur de ϕ .

Exemple 13 Soit le changement de variable $\phi = au^2$. Dans cet exemple, les valeurs négatives et positives de u donnent la même valeur de ϕ . On aurait dans ce cas que $\mathcal{W}(\phi) = 2\mathcal{P}(u) \left| \frac{du}{d\phi} \right|$ puisque deux intervalles disjoints du donnent le même intervalle $d\phi$. Ceci est représenté graphiquement à la Fig.(1-10)

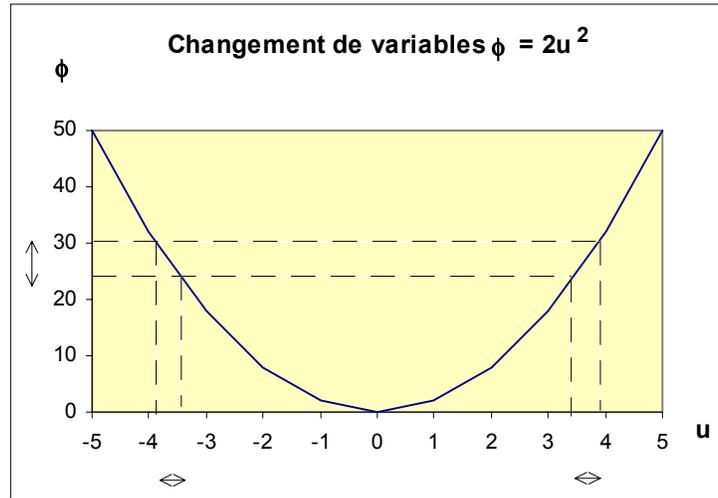


Figure 1-10 Changement de variables $\phi = 2u^2$

Dans l'exemple précédent, il y a deux façons différentes d'obtenir le même ϕ . Le facteur 2 reflète donc la règle du OU pour des événements complémentaires. De façon plus générale, lorsque la transformation entre ϕ et u n'est pas biunivoque, il faut tenir compte de toutes les valeurs de u qui donnent le même ϕ .

Remarque 17 Le résultat pour une transformation générale $\phi(u)$ est que s'il y a n valeurs de u , notées u_i ($i = 1, 2 \dots n$), correspondant à une valeur de ϕ donnée, alors

$$\mathcal{W}(\phi) = \sum_{i=1}^n \mathcal{P}(u_i) \left| \frac{du}{d\phi} \right|_{u_i} \quad (1.98)$$

Voici maintenant un exemple, tiré textuellement de Reif, qui contient les deux difficultés mentionnées ci-dessus.

Exemple 14 Soit un vecteur bidimensionnel \mathbf{r} de longueur fixe mais d'orientation aléatoire. La probabilité à priori que le vecteur pointe dans la direction θ est donnée par

$$\mathcal{P}(\theta) d\theta = \frac{d\theta}{2\pi} \quad (1.99)$$

Quelle est la probabilité que la composante

$$x = |\mathbf{r}| \cos \theta$$

du vecteur prenne une valeur donnée? Ce problème est illustré à la figure (1-11). Notons d'abord que les valeurs négatives et positives de θ donnent la même valeur de x . On a donc,

$$\mathcal{W}(x) = 2\mathcal{P}(\theta) \left| \frac{d\theta}{dx} \right| = \frac{1}{\pi} \left| \frac{d\theta}{dx} \right| \quad (1.100)$$

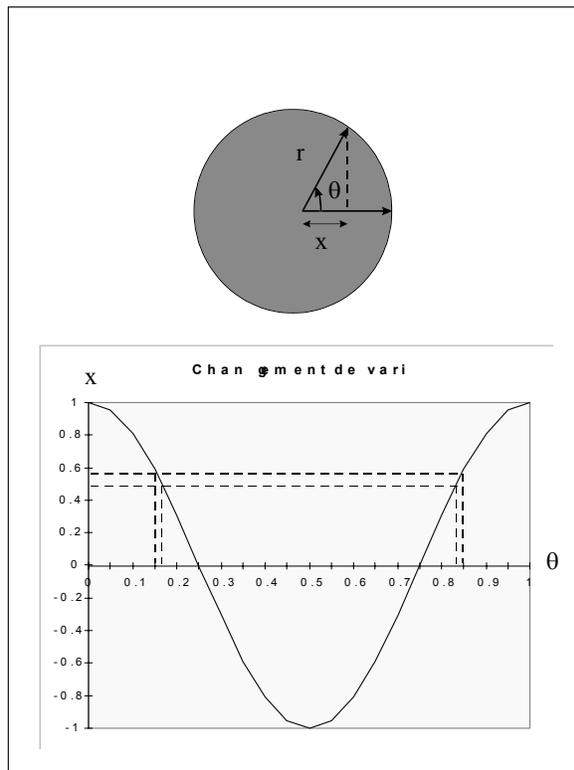


Figure 1-11 Changement de variables, d'un angle à la projection du vecteur.

et on peut calculer la dérivée comme d'habitude

$$\left| \frac{dx}{d\theta} \right| = |\mathbf{r}| |\sin \theta| = |\mathbf{r}| \sqrt{1 - \cos^2 \theta} = |\mathbf{r}| \sqrt{1 - \frac{x^2}{|\mathbf{r}|^2}} \quad (1.101)$$

de telle sorte que finalement, comme $|d\theta/dx| = |dx/d\theta|^{-1}$, alors

$$\mathcal{W}(x) = \begin{cases} \frac{1}{\pi} \frac{1}{\sqrt{|\mathbf{r}|^2 - x^2}} & ; \text{ pour } -|\mathbf{r}| \leq x \leq |\mathbf{r}| \\ 0 & ; \text{ autrement} \end{cases} \quad (1.102)$$

Notez que, contrairement à une probabilité, une "densité" de probabilité peut diverger, ce qui est le cas ici pour $x = \pm |\mathbf{r}|$. Tant que la divergence est intégrable, il n'y a évidemment pas de problème.

Remarque 18 En terminant, on peut discuter brièvement le cas de plusieurs variables de façon tout à fait analogue. Soit une densité de probabilité conjointe pour deux variables, c'est-à-dire que la probabilité de trouver à la fois x et y dans l'intervalle $dxdy$ est donnée par $\mathcal{P}(x, y) dxdy$. Alors, si on fait le changement de variables $u(x, y)$ et $v(x, y)$, la densité de probabilité pour les variables u et v se trouve en utilisant la formule de changement de variable pour les intégrales, soit

$$\mathcal{W}(u, v) = \mathcal{P}(x, y) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| \quad (1.103)$$

où le jacobien $\frac{\partial(x, y)}{\partial(u, v)}$ comme d'habitude est défini par

$$\frac{\partial(x, y)}{\partial(u, v)} = \det \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix} \quad (1.104)$$

Les problèmes reliés au fait que la transformation n'est peut-être pas biunivoque sont réglés de la même façon que ci-haut.

Remarque 19 *Juste pour se rappeler un peu d'où vient le Jacobien lorsqu'on transforme des surfaces, considérons le cas d'une transformation $x(u, v)$, $y(u, v)$. Soit le vecteur infinitésimal de composantes du_1 et dv_1 partant du point (u, v) . Le vecteur correspondant en (x, y) a les composantes

$$dx_1 = \frac{\partial x}{\partial u} du_1 + \frac{\partial x}{\partial v} dv_1 \quad (1.105)$$

$$dy_1 = \frac{\partial y}{\partial u} du_1 + \frac{\partial y}{\partial v} dv_1 \quad (1.106)$$

Un autre vecteur du_2 et dv_2 partant du même point (u, v) aurait des composantes dx_2 , dy_2 données par une formule analogue. Rappelons-nous maintenant qu'en deux dimensions, la valeur absolue de la surface engendrée par les deux vecteurs de composantes dx_1 , dy_1 et dx_2 , dy_2 est donnée par le produit vectoriel de ces deux vecteurs. Ce produit vectoriel n'est autre que le déterminant. Donc, utilisant la formule précédente pour le changement de variable on obtient

$$\left| \det \begin{bmatrix} dx_1 & dx_2 \\ dy_1 & dy_2 \end{bmatrix} \right| = \left| \det \left\{ \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix} \begin{bmatrix} du_1 & du_2 \\ dv_1 & dv_2 \end{bmatrix} \right\} \right| \quad (1.107)$$

Puis, utilisant la propriété $\det(AB) = \det(A)\det(B)$ on voit que le jacobien donne la constante de proportionnalité entre les deux surfaces. De même, pour trois variables le produit triple de trois vecteurs (le volume) est un déterminant et, en suivant le même raisonnement, on voit que le jacobien représente la constante de proportionnalité entre les deux volumes.

1.6 Discussion générale de la marche aléatoire et loi des grands nombres

D'après notre discussion de la gaussienne, on se doute bien que c'est la distribution de probabilité qu'on doit obtenir, indépendamment de bien des détails de la probabilité à priori pour les pas à gauche et à droite. Par exemple, jusqu'à maintenant, on a supposé qu'on ne peut faire que des pas à gauche et à droite tous de la même longueur ℓ avec une probabilité respective p et q . La distribution de probabilité pour la longueur d'un pas quelconque, dans le cas $\ell = 1$, est représentée en haut à gauche de la figure (1-12).

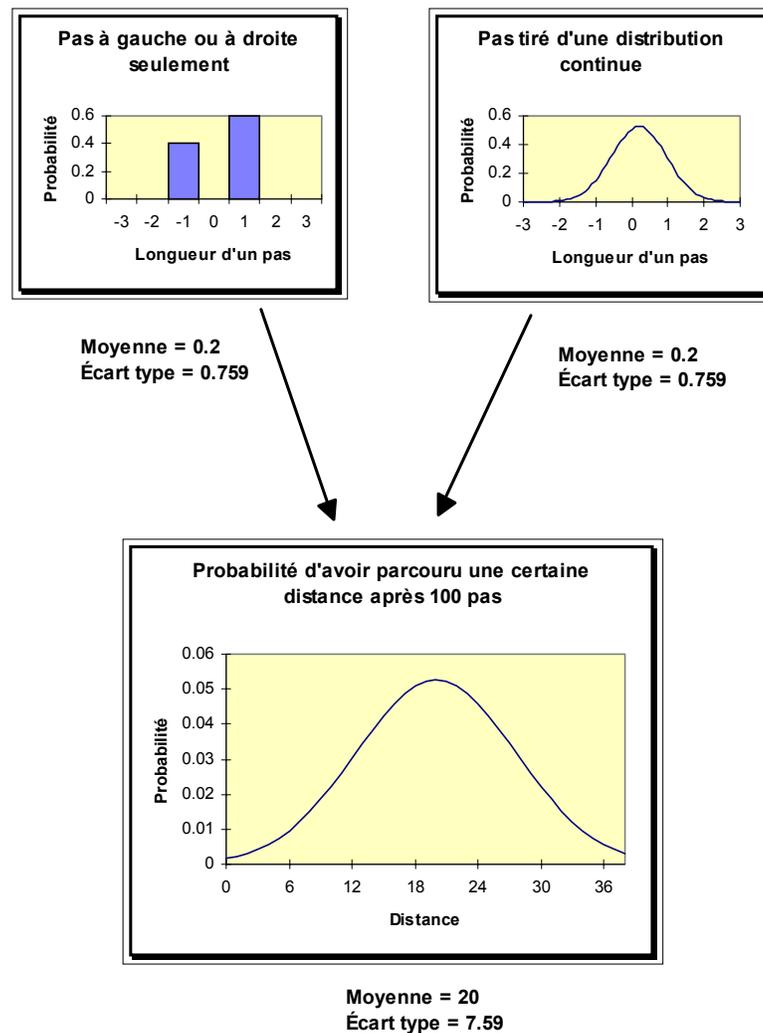


Figure 1-12 Illustration du théorème de la limite centrale.

Il est remarquable qu'une distribution de départ très asymétrique pour un seul pas, comme celle d'en haut à gauche de la figure, donne une distribution finale pour la distance totale parcourue qui soit lisse et symétrique comme dans le bas de la figure. Dans une section précédente nous avons prouvé que cette distribution était une gaussienne. On peut déjà deviner que ce résultat est très général. Les

détails de la distribution de probabilité de départ sont “oubliés”. Soyons plus spécifiques. Supposons que la grandeur et la direction du pas (sa valeur algébrique) soient plutôt tirés d’une distribution de probabilité à priori qui est continue tel qu’illustré en haut à droite de la figure. À condition que ces deux lois de probabilité très différentes (celles de gauche et de droite en haut de la figure) aient la même valeur moyenne et le même écart type, après un grand nombre N de pas, la probabilité d’être à une certaine distance est donnée par la même gaussienne, avec une valeur moyenne qui est égale à N fois la grandeur moyenne du pas obtenu de la probabilité à priori pour un seul pas et un écart-type égal à \sqrt{N} fois l’écart type de la probabilité à priori. Dans l’exemple de la figure, $N = 100$.

Nous ne nous lancerons pas dans la preuve que la distribution de probabilité est gaussienne dans la limite $N \rightarrow \infty$ (sauf pour donner à la section 1.8 une preuve de ce “théorème de la limite centrale” reposant sur des notions avancées) mais, par contre, nous allons prouver les relations entre les écarts types de la distribution de probabilité à priori pour une variable, disons s , et celle de la distribution de probabilité pour la somme $\sum_{i=1}^N s_i$, où s_i est la valeur prise par la variable aléatoire s au i ème tirage. Ces résultats ont des applications très importantes dans l’analyse de l’erreur sur la moyenne et le problème de l’élimination du bruit dans les signaux, sans mentionner les sondages!

1.6.1 Distribution de probabilité pour plusieurs variables¹⁶

Cette sous-section sert à présenter les notions de probabilité pour plusieurs variables. Elle est donc en quelque sorte un lemme pour la sous-section suivante. Nous ferons la discussion dans le cas de variables continues, mais les résultats se transposent facilement au cas des variables discrètes.

Soit $\mathcal{P}(x, y) dx dy$ la *probabilité conjointe* que les deux variables aléatoires prennent respectivement les valeurs x et y dans l’intervalle $dx dy$. La valeur moyenne d’une fonction quelconque $f(x, y)$ de ces variables aléatoires se calcule ainsi

$$\langle f \rangle = \bar{f} = \int dx \int dy \mathcal{P}(x, y) f(x, y) \quad (1.108)$$

Évidemment, cette définition mène immédiatement au résultat suivant

$$\boxed{\langle f + g \rangle = \langle f \rangle + \langle g \rangle} \quad (1.109)$$

c’est-à-dire que l’opérateur “moyenne” $\langle \rangle$ est un opérateur linéaire.

Cas des variables statistiquement indépendantes: Par définition, les variables x et y sont statistiquement indépendantes si la probabilité d’obtenir x n’influence pas la probabilité d’obtenir y et vice-versa. Dans ce cas, notre discussion générale sur la façon de combiner les probabilités, telle que donnée au début de ce chapitre, nous dit que la probabilité d’obtenir x et y est simplement égale au produit des probabilités, c’est-à-dire,

$$\boxed{\mathcal{P}(x, y) dx dy = P_1(x) dx P_2(y) dy \quad ; \quad \text{pour des variables stat. indep.}} \quad (1.110)$$

Pour des variables statistiquement indépendantes, on a donc

$$\boxed{\langle f(x) g(y) \rangle = \langle f(x) \rangle \langle g(y) \rangle} \quad (1.111)$$

¹⁶Reif, Sec.1.7

La preuve est donnée dans les quelques lignes suivantes (qui ne sont qu'une façon d'écrire la distributivité du produit sur la somme du point de vue des intégrales)

$$\langle f(x)g(y) \rangle = \int \int dx dy P_1(x) P_2(y) f(x) g(y) \quad (1.112)$$

$$= \int dx P_1(x) f(x) \int dy P_2(y) g(y) \quad (1.113)$$

$$= \langle f(x) \rangle \langle g(y) \rangle \quad (1.114)$$

Lorsque des variables ne sont pas statistiquement indépendantes, on utilise généralement $\langle f(x)g(y) \rangle - \langle f(x) \rangle \langle g(y) \rangle$ comme une mesure des *corrélations*.

1.6.2 Discussion générale des valeurs moyennes pour la marche aléatoire¹⁷

On peut décrire la marche aléatoire de la façon suivante. Soit s une variable aléatoire qui vaut 1 si on fait un pas à droite, et 0 si on fait un pas à gauche. Cette variable aléatoire peut être utilisée pour compter le nombre de pas faits à droite. Soient p la probabilité à priori pour que $s = 1$ et $q = 1 - p$ la probabilité à priori pour que $s = 0$. Commençons par considérer la valeur moyenne et l'écart type pour cette variable aléatoire.

$$\langle s \rangle = 1 \cdot p + 0 \cdot q = p \quad (1.115)$$

$$\langle s^2 \rangle - \langle s \rangle^2 = (1 \cdot p + 0 \cdot q) - p^2 = p(1 - p) = pq \quad (1.116)$$

Chaque pas de la marche aléatoire est décrit par une telle variable s_i et la distribution de probabilité pour chacune de ces variables s_i a la même forme fonctionnelle.¹⁸ Chaque pas est de plus statistiquement indépendant des autres. Le nombre total de pas à droite dans une marche quelconque est donné par

$$n = \sum_{i=1}^N s_i \quad (1.117)$$

À partir des résultats de la section précédente, on obtient donc nos résultats fondamentaux.

Moyenne:

$$\langle n \rangle = \left\langle \sum_{i=1}^N s_i \right\rangle = \sum_{i=1}^N \langle s_i \rangle = N \langle s \rangle \quad (1.118)$$

Ce résultat est indépendant de la forme particulière de notre distribution de probabilité pour s . Dans le cas particulier considéré au début de cette sous-section, on a donc

$$\langle n \rangle = Np \quad (1.119)$$

ce qui est bien ce que nous avons obtenu avec la binomiale.

Écart type:

$$\langle n^2 \rangle - \langle n \rangle^2 = N \left(\langle s^2 \rangle - \langle s \rangle^2 \right) \quad (1.120)$$

¹⁷Reif, Sec. 1.9

¹⁸En mathématiques, ceci s'appelle des essais de Bernoulli. Des essais répétés, statistiquement indépendants, s'appellent des "essais de Bernoulli" s'il n'y a que deux résultats possibles pour chaque essai et si les probabilités sont les mêmes pour tous les essais. (Nommé en l'honneur de James Bernoulli, 1654-1705).

un résultat qu'on prouve facilement de la façon suivante.

$$\langle n^2 \rangle - \langle n \rangle^2 = \left\langle \sum_{i=1}^N \sum_{j=1}^N s_i s_j \right\rangle - \left\langle \sum_{i=1}^N s_i \right\rangle^2 \quad (1.121)$$

$$= \left\langle \sum_{i=1}^N s_i^2 \right\rangle + \left\langle \sum_{i=1}^N \sum_{j \neq i=1}^N s_i s_j \right\rangle - \left\langle \sum_{i=1}^N s_i \right\rangle^2 \quad (1.122)$$

$$= N \langle s^2 \rangle + N(N-1) \langle s \rangle \langle s \rangle - N^2 \langle s \rangle^2 \quad (1.123)$$

$$= N \left(\langle s^2 \rangle - \langle s \rangle^2 \right) \quad (1.124)$$

Le seul point un peu subtil de ces manipulations est celui de la double somme. Lorsque $i = j$ nous devons considérer $\langle s_i^2 \rangle$ mais lorsque $i \neq j$, nous avons plutôt $\langle s_i s_j \rangle = \langle s_i \rangle \langle s_j \rangle$. Comme ci-haut, ce résultat est indépendant de la distribution de probabilité alors pour le cas particulier considéré en début de sous-section, nous avons

$$\langle n^2 \rangle - \langle n \rangle^2 = N \left(\langle s^2 \rangle - \langle s \rangle^2 \right) = Npq \quad (1.125)$$

tel que nous avons trouvé avec la binomiale.

Pour une variable qui est la somme de N variables statistiquement indépendants de moyenne $\langle s \rangle$ et d'écart type $\langle s^2 \rangle$ nous avons donc

$$\boxed{\frac{\sqrt{\langle n^2 \rangle - \langle n \rangle^2}}{\langle n \rangle} = \frac{1}{\sqrt{N}} \frac{\sqrt{\langle s^2 \rangle - \langle s \rangle^2}}{\langle s \rangle}} \quad (1.126)$$

un résultat fondamental. Le résultat encore plus général est que la distribution de probabilité de n est une gaussienne. C'est le *théorème de la limite centrale*. Ou la *loi des grands nombres*.

Remarque 20 *Évidemment, il faut que la moyenne et l'écart type de la distribution de probabilité de départ, ici celle des s , existe pour que le théorème de la limite centrale s'applique. Il y a des distributions de probabilité dont ni la moyenne ni l'écart type n'existent. Un exemple d'une telle distribution de probabilité est donné par $P(x) = Ax^{-3/2}\theta(x-a)$ où A est une constante assurant la normalisation et θ est la fonction de Heaviside,

$$\theta(x) \equiv \begin{cases} 1 & ; \quad x > 0 \\ 0 & ; \quad x < 0 \end{cases} \quad (1.127)$$

On peut généraliser alors le théorème de la limite centrale mais les distributions de probabilité qu'on obtient sont des distributions dites de Lévy. (cf. *La Recherche*, juillet 1997). Ces distributions sont intimement reliées à la notion de fractale.

Exemple 15 Le résultat Éq.(1.126) sur le rétrécissement relatif de la distribution de probabilité lorsqu'une variable est la somme de N variables indépendantes a des applications importantes en théorie de la mesure expérimentale. En effet, si on fait une moyenne sur N mesures expérimentales entachées de bruit, alors la valeur moyenne des mesures est décrite par une distribution de probabilité dont l'écart type est \sqrt{N} fois plus petit que l'écart type du bruit qui entache les mesures, en autant que le bruit ne soit pas corrélé d'une mesure à l'autre. En effet, soit la moyenne expérimentale x de N mesures m_i

$$x = \frac{\sum_{i=1}^N m_i}{N} \quad (1.128)$$

alors

$$\langle x \rangle = N \left\langle \frac{m}{N} \right\rangle = \langle m \rangle \quad (1.129)$$

et

$$\langle x^2 \rangle - \langle x \rangle^2 = N \left(\left\langle \left(\frac{m}{N} \right)^2 \right\rangle - \left\langle \frac{m}{N} \right\rangle^2 \right) = \frac{1}{N} (\langle m^2 \rangle - \langle m \rangle^2) \quad (1.130)$$

En d'autres mots, d'un point de vue statistique, l'erreur attendue sur la variable x qui représente la moyenne expérimentale Eq.(1.128) est \sqrt{N} fois plus petite que l'erreur qu'on observerait sur chacune des mesures m_i prises individuellement.

1.7 *Formule de Stirling

La formule de Stirling¹⁹ a été utilisée dans ce chapitre pour démontrer que la binomiale devient une gaussienne dans la limite N grand. La démonstration de la formule de Stirling illustre une méthode d'approximation très générale qui est souvent utilisée en physique statistique, l'approximation du col. On retrouve cette méthode d'approximation par exemple dans la preuve des équivalences entre ensembles, dans la preuve du théorème de la limite centrale et dans la théorie des transitions de phase. C'est donc en partie la généralité de la méthode de dérivation qui fait que nous nous y attardons. Dans le domaine général des mathématiques, l'approximation du col est souvent utilisée pour dériver des séries asymptotiques.

On commence par représenter $N!$ par une intégrale. La représentation intégrale que nous cherchons, celle de la fonction gamma Γ d'Euler est la suivante

$$\Gamma(N+1) \equiv N! = \int_0^\infty dx x^N e^{-x} \quad (1.131)$$

La preuve de ce résultat s'obtient simplement en intégrant par parties

$$\int_0^\infty dx x^N e^{-x} = -x^N e^{-x} \Big|_0^\infty + N \int_0^\infty dx x^{N-1} e^{-x} \quad (1.132)$$

$$= N \int_0^\infty dx x^{N-1} e^{-x} \quad (1.133)$$

Continuant ainsi jusqu'à

$$\int_0^\infty dx e^{-x} = 1 \quad (1.134)$$

prouve le résultat.

Pour approximer l'intégrale dans la représentation intégrale Eq.(1.131) on la réécrit sous la forme

$$\int_0^\infty dx x^N e^{-x} = \int_0^\infty dx e^{N \ln x} e^{-x} = \int_0^\infty dx e^{N \ln x - x} \quad (1.135)$$

et on note que l'argument de l'exponentielle a un maximum très prononcé. Ce maximum est situé à

$$\frac{d}{dx} (N \ln x - x) = \frac{N}{x} - 1 = 0 \quad (1.136)$$

Faisant un développement limité (série de Taylor) autour du maximum à $x = N$ on trouve

$$N \ln x - x = (N \ln N - N) - \frac{1}{2N} (x - N)^2 + \left(\frac{1}{3N^2} \right) (x - N)^3 + \mathcal{O}((x - N)^4) \quad (1.137)$$

¹⁹Reif, annexe A.6

Substituant dans l'intégrale que nous cherchons à évaluer, nous obtenons

$$\begin{aligned} N! &= \int_0^\infty dx e^{N \ln x - x} & (1.138) \\ &\simeq \int_0^\infty dx \exp \left[(N \ln N - N) - \frac{1}{2N} (x - N)^2 + \left(\frac{1}{3N^2} \right) (x - N)^3 + \dots \right] \end{aligned}$$

L'exponentielle chute de $\exp(-1/2)$ aussitôt que x s'éloigne du maximum de \sqrt{N} . Pour $x = N + \sqrt{N}$ donc, on trouve que le terme suivant $(x - N)^3 / (3N^2)$ dans l'argument de l'exponentielle remplace $\exp(-1/2)$ par $\exp(-1/2 + \mathcal{O}(1/\sqrt{N}))$. On néglige donc cette correction et il nous reste

$$N! \simeq \exp[N \ln N - N] \int_0^\infty dx \exp \left[-\frac{1}{2N} (x - N)^2 \right] \quad (1.139)$$

En notant qu'on peut étendre la borne inférieure de l'intégrale jusqu'à $-\infty$ avec seulement une correction d'ordre e^{-N} , on a

$$N! \simeq \exp[N \ln N - N] \int_{-\infty}^\infty dx \exp \left[-\frac{1}{2N} (x - N)^2 \right] \quad (1.140)$$

qu'on évalue en utilisant nos résultats pour les intégrales gaussiennes

$$N! \simeq \sqrt{2\pi N} \exp[N \ln N - N] \quad (1.141)$$

ce qui nous redonne bien la formule de Stirling Éq.(1.7). En tenant compte des corrections suivantes, Reif²⁰ donne la première correction,

$$N! \simeq \sqrt{2\pi N} \exp[N \ln N - N] \left[1 + \frac{1}{12N} + \dots \right]. \quad (1.142)$$

1.8 *Théorème de la limite centrale

La démonstration donnée dans cette annexe est quelque peu différente de celle de Reif.²¹ Vous pouvez donc choisir celle que vous préférez.

On commence par introduire une définition. La fonction caractéristique d'une distribution de probabilité continue est simplement sa transformée de Fourier

$$Q(k) \equiv \langle e^{-ikx} \rangle \equiv \int dx P(x) e^{-ikx} \quad (1.143)$$

Si on connaît la fonction caractéristique d'une distribution de probabilité, alors on obtient évidemment la distribution de probabilité en prenant la transformée de Fourier inverse de la distribution de probabilité. De plus, on peut calculer tous les moments de la distribution de probabilité de façon très simple puisque

$$\left(i \frac{\partial}{\partial k} \right)^n \langle e^{-ikx} \rangle \Big|_{k=0} = \langle x^n \rangle \quad (1.144)$$

On appelle donc aussi *fonction génératrice des moments* cette fonction caractéristique.

²⁰Reif, p.614

²¹Reif, Sec. 1.11

Si une variable aléatoire X est la somme de N variables aléatoires indépendantes distribuées de façon identique (i.e. chaque variable aléatoire a la même distribution de probabilité) alors la variable X

$$X = \sum_{i=1}^N x_i \quad (1.145)$$

a une distribution de probabilité qui est gaussienne. Pour prouver ce *théorème de la limite centrale*, nous allons utiliser une variable un peu différente, soit

$$Y = \sqrt{N} \left[\left(\frac{1}{N} \sum_{i=1}^N x_i \right) - \mu \right] \quad (1.146)$$

et montrer que dans la limite $N \rightarrow \infty$, la distribution de probabilité pour la variable Y a une limite unique si μ est choisi convenablement, mais de façon indépendante de N . Ce qui motive le choix de Y est simplement les graphiques dont nous nous sommes servis pour illustrer la limite $N \rightarrow \infty$ de la binomiale. Dans ce cas, nous avons choisi l'origine et l'échelle pour que la distribution de probabilité tombe toujours dans notre champ de vision. Nous pouvons toujours retrouver la distribution de probabilité pour la variable X à l'aide d'un changement de variable simple, soit

$$X = \left(\frac{Y}{\sqrt{N}} + \mu \right) N \quad (1.147)$$

Pour prouver qu'effectivement il existe une limite pour la fonction Y , il suffit de prouver que sa fonction caractéristique a une limite, qu'on calculera explicitement

$$\begin{aligned} \lim_{N \rightarrow \infty} \langle e^{-ikY} \rangle &= \lim_{N \rightarrow \infty} \left\langle e^{-ik\sqrt{N} \left[\left(\frac{1}{N} \sum_{i=1}^N x_i \right) - \mu \right]} \right\rangle = \\ &= \lim_{N \rightarrow \infty} \left\langle e^{ik\sqrt{N}\mu} \prod_{i=1}^N e^{-ikx_i/\sqrt{N}} \right\rangle \\ &= \lim_{N \rightarrow \infty} e^{ik\sqrt{N}\mu} \left\langle e^{-ikx/\sqrt{N}} \right\rangle^N \end{aligned} \quad (1.148)$$

où dans la dernière égalité nous avons utilisé le théorème sur la moyenne d'un produit de fonctions de variables aléatoires indépendantes Éq.(1.111).

Comme $N \rightarrow \infty$, on peut développer $\left\langle e^{-ikx/\sqrt{N}} \right\rangle$ en puissances de ikx/\sqrt{N} . Cependant, nous faisons face à une difficulté déjà rencontrée, soit celle que $\left\langle e^{-ikx/\sqrt{N}} \right\rangle$ est élevée à la puissance N avec $N \rightarrow \infty$. Donc, si nous écrivons

$$\left\langle e^{-ikx/\sqrt{N}} \right\rangle^N \simeq \left\langle 1 - \frac{ikx}{\sqrt{N}} + \dots \right\rangle^N \simeq 1 - \frac{ikN \langle x \rangle}{\sqrt{N}} + \dots \quad (1.149)$$

la première correction est immense, $ik\sqrt{N} \langle x \rangle$, plutôt que d'être petite! Il est donc préférable de passer par l'artifice suivant. Supposons qu'on parvienne à écrire

$$\left\langle e^{-ikx/\sqrt{N}} \right\rangle = e^{f(k/\sqrt{N})} \quad (1.150)$$

où $f(k/\sqrt{N})$ est un fonction de k/\sqrt{N} qu'on peut trouver simplement en prenant le logarithme de $\left\langle e^{-ikx/\sqrt{N}} \right\rangle$. Alors, nous aurons

$$\left\langle e^{-ikx/\sqrt{N}} \right\rangle^N = e^{Nf(k/\sqrt{N})} \quad (1.151)$$

et le développement de $f(k/\sqrt{N})$ en puissances de k/\sqrt{N} sera cette fois-ci permis.

Remarque 21 Soit

$$f(k) = \ln \langle e^{-ikx} \rangle \quad (1.152)$$

Cette fonction $f(k)$ est la fonction génératrice des cumulants de la distribution de probabilité, c'est-à-dire qu'on définit le cumulant $\langle x^n \rangle_c$ d'ordre n par la relation

$$\left(i \frac{\partial}{\partial k} \right)^n f(k) \equiv \langle x^n \rangle_c \quad (1.153)$$

Dans le cas qui nous intéresse l'argument de f est $k/\sqrt{N} \ll 1$. Le développement limité nous donne,

$$f\left(\frac{k}{\sqrt{N}}\right) = \ln \langle e^{-ikx/\sqrt{N}} \rangle \simeq \ln \left\langle \left(1 - i \frac{k}{\sqrt{N}} x - \frac{1}{2} \frac{k^2}{N} x^2 + \mathcal{O}\left(\frac{k^3}{N^{3/2}}\right) \right) \right\rangle \quad (1.154)$$

en utilisant $\ln(1+y) = y - \frac{1}{2}y^2 + \dots$ on a ensuite

$$f\left(\frac{k}{\sqrt{N}}\right) \simeq -i \frac{k}{\sqrt{N}} \langle x \rangle - \frac{1}{2} \frac{k^2}{N} (\langle x^2 \rangle - \langle x \rangle^2) + \mathcal{O}\left(\frac{k^3}{N^{3/2}}\right) \quad (1.155)$$

Substituant ce résultat dans Éq.(1.151) et dans l'expression pour la fonction caractéristique de Y nous obtenons,

$$\lim_{N \rightarrow \infty} \langle e^{-ikY} \rangle = \lim_{N \rightarrow \infty} e^{ik\sqrt{N}\mu} \langle e^{-ikx/\sqrt{N}} \rangle^N \quad (1.156)$$

$$= \lim_{N \rightarrow \infty} e^{ik\sqrt{N}\mu} e^{Nf(k/\sqrt{N})} \quad (1.157)$$

$$= \lim_{N \rightarrow \infty} e^{ik\sqrt{N}\mu} e^{-ik\sqrt{N}\langle x \rangle - \frac{1}{2}k^2(\langle x^2 \rangle - \langle x \rangle^2) + \mathcal{O}\left(\frac{k^3}{N^{1/2}}\right)} \quad (1.158)$$

$$= \lim_{N \rightarrow \infty} e^{ik\sqrt{N}\mu - ik\sqrt{N}\langle x \rangle - \frac{1}{2}k^2(\langle x^2 \rangle - \langle x \rangle^2)} \quad (1.159)$$

$$= e^{-\frac{1}{2}k^2(\langle x^2 \rangle - \langle x \rangle^2)} \quad (1.160)$$

où la dernière ligne est la limite que nous cherchions et qui n'existe que si on fait le choix

$$\mu = \langle x \rangle \quad (1.161)$$

À partir de la fonction caractéristique pour Y , on obtient

$$\langle Y \rangle = \left(i \frac{\partial}{\partial k} \right) \langle e^{-ikY} \rangle \Big|_{k=0} = \left(i \frac{\partial}{\partial k} \right) e^{-\frac{1}{2}k^2(\langle x^2 \rangle - \langle x \rangle^2)} \Big|_{k=0} = 0 \quad (1.162)$$

$$\langle Y^2 \rangle = \left(i \frac{\partial}{\partial k} \right)^2 e^{-\frac{1}{2}k^2(\langle x^2 \rangle - \langle x \rangle^2)} \Big|_{k=0} = \langle x^2 \rangle - \langle x \rangle^2 \quad (1.163)$$

De ces résultats, on trouve facilement la moyenne et l'écart type pour la variable de départ X , à partir de sa relation à Y , Éq.(1.147)

$$\langle X \rangle = \left(\frac{\langle Y \rangle}{\sqrt{N}} + \mu \right) N = N\mu = N \langle x \rangle \quad (1.164)$$

comme on s'y attendait, et

$$\langle (X - \langle X \rangle)^2 \rangle = N \langle Y^2 \rangle = N (\langle x^2 \rangle - \langle x \rangle^2). \quad (1.165)$$

Ou encore, on peut trouver directement la fonction caractéristique pour la variable X , soit

$$\langle e^{-ikX} \rangle = e^{-ikN\langle x \rangle} \langle e^{-ik\sqrt{N}Y} \rangle \quad (1.166)$$

$$= e^{-ikN\langle x \rangle} e^{-\frac{1}{2}k^2 N(\langle x^2 \rangle - \langle x \rangle^2)} \quad (1.167)$$

$$= e^{-ik\mu'} e^{-\frac{1}{2}k^2 \sigma'^2} \quad (1.168)$$

et de là évaluer toutes les moyennes. Ci-dessus, nous avons défini

$$\mu' = N\langle x \rangle \quad ; \quad \sigma'^2 = N(\langle x^2 \rangle - \langle x \rangle^2). \quad (1.169)$$

La distribution de probabilité pour X elle-même s'évalue en prenant la transformée de Fourier. Or, la transformée de Fourier d'une gaussienne est une gaussienne. En effet, il suffit de compléter le carré dans l'argument pour prouver ce résultat

$$\mathcal{P}(X) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{ikX} e^{-ik\mu'} e^{-\frac{1}{2}k^2 \sigma'^2} \quad (1.170)$$

$$= \int_{-\infty}^{\infty} \frac{dk}{2\pi} \exp \left[-\frac{1}{2}\sigma'^2 \left(k - \frac{iX - i\mu'}{\sigma'^2} \right)^2 - \frac{(X - \mu')^2}{2\sigma'^2} \right] \quad (1.171)$$

$$= \exp \left[-\frac{(X - \mu')^2}{2\sigma'^2} \right] \int_{-\infty}^{\infty} \frac{dk}{2\pi} \exp \left[-\frac{1}{2}\sigma'^2 \left(k - \frac{iX - i\mu'}{\sigma'^2} \right)^2 \right] \quad (1.172)$$

$$= \frac{1}{\sqrt{2\pi\sigma'^2}} \exp \left[-\frac{(X - \mu')^2}{2\sigma'^2} \right] \quad (1.173)$$

Remarque 22 Pour une gaussienne, tous les cumulants d'ordre plus élevé que le deuxième s'annulent.

1.9 *Chi carré et estimé de l'écart type

Dans la pratique, il est souvent intéressant de pouvoir estimer la précision avec laquelle l'écart-type a été obtenu d'un nombre limité de mesures. Supposons que l'on fasse une seule mesure dont la densité de probabilité est donnée par une loi normale. Soit donc $y = x^2$. En utilisant notre formule de changement de variables on obtient

$$P(y) = \mathcal{P}(x) \frac{dx}{dy} \quad (1.174)$$

où nous allons supposer que $\mathcal{P}(x)$ est une gaussienne de moyenne $\mu = 0$ et d'écart type σ . (Le cas $\mu \neq 0$ s'obtient par une translation triviale). On a donc,

$$\mathcal{P}(x) \frac{dx}{dy} = \frac{1}{2x} e^{\left(-\frac{x^2}{2\sigma^2}\right)} = P(y) = \frac{1}{2\sqrt{y}} e^{\left(-\frac{y}{2\sigma^2}\right)} \quad (1.175)$$

La dernière expression est $f_{\frac{1}{2\sigma^2}, \frac{1}{2}}(y)$, un cas particulier de ce qui s'appelle la densité de probabilité gamma définie par l'expression générale²²

$$f_{\alpha, \nu}(y) = \frac{1}{\Gamma(\nu)} \alpha^\nu y^{\nu-1} e^{-\alpha y} \quad (1.176)$$

²²Dans cette section nous utilisons les définitions et les preuves décrites dans W. Feller, Volume II, (Wiley, New York, 1971) pp.47,48.

Le paramètre α joue le rôle d'un facteur d'échelle.

Remarque 23 Cette densité de probabilité est normalisée puisque la fonction gamma d'Euler est définie par

$$\Gamma(\nu) = \int_0^{\infty} x^{\nu-1} e^{-x} dx \quad (1.177)$$

La fonction gamma est une généralisation aux réels des factorielles définies sur les entiers. En effet, il est facile de démontrer en intégrant par parties que pour n entier,

$$\Gamma(n) = (n-1)! \quad (1.178)$$

Le cas qui nous intéresse est celui où on fait N mesures indépendantes. Supposons qu'on veuille connaître la densité de probabilité pour la variable²³

$$\chi_N^2 = \sum_{i=1}^N x_i^2 \quad (1.179)$$

On a trouvé ci-dessus que la densité de probabilité pour $N = 1$ est $f_{\frac{1}{2\sigma^2}, \frac{1}{2}}(\chi_1^2)$. On obtient la densité de probabilité pour $N = 2$ en multipliant la probabilité pour obtenir $\chi_1^2 (< \chi_2^2)$ pour la première mesure et $\chi_2^2 - \chi_1^2$ pour la deuxième mesure, puis en sommant sur toutes les valeurs possibles de χ_1^2 . Mathématiquement, cette procédure mène au résultat,

$$P_2(\chi_2^2) = \int_0^{\chi_2^2} dy f_{\frac{1}{2\sigma^2}, \frac{1}{2}}(\chi_2^2 - \chi_1^2) f_{\frac{1}{2\sigma^2}, \frac{1}{2}}(\chi_1^2) \quad (1.180)$$

qui n'est rien d'autre que la convolution de densités de probabilité qui s'annulent pour des arguments négatifs. Or, tout comme la convolution de deux gaussiennes donne une gaussienne, la convolution de deux densités gamma de même paramètre d'échelle α donne une densité gamma. Plus précisément,

$$\boxed{f_{\alpha, \nu+\mu}(y) = \int_0^y dx f_{\alpha, \nu}(y-x) f_{\alpha, \mu}(x)} \quad (1.181)$$

Preuve:

$$\int_0^y dx f_{\alpha, \nu}(y-x) f_{\alpha, \mu}(x) = \frac{\alpha^\nu \alpha^\mu}{\Gamma(\nu) \Gamma(\mu)} e^{-\alpha y} \int_0^y dx (y-x)^{\nu-1} x^{\mu-1} \quad (1.182)$$

Le changement de variable $x = ty$ donne,

$$= \frac{\alpha^{\nu+\mu}}{\Gamma(\nu) \Gamma(\mu)} e^{-\alpha y} x^{\nu+\mu-1} \int_0^1 dt (1-t)^{\nu-1} t^{\mu-1} \quad (1.183)$$

$$= f_{\alpha, \nu+\mu}(y) \quad (1.184)$$

La constante de normalisation n'a pas besoin d'être explicitement vérifiée puisque la convolution préserve la normalisation. Néanmoins, on sait que

$$\int_0^1 dt (1-t)^{\nu-1} t^{\mu-1} \equiv B(\nu, \mu) = \frac{\Gamma(\nu) \Gamma(\mu)}{\Gamma(\nu + \mu)} \quad (1.185)$$

²³Il y a différentes définitions du χ^2 mais elle sont en général trivialement reliées par un changement d'échelle.

Utilisant le résultat précédent sur la convolution, on a pour N mesures que

$$\boxed{P_N(\chi_N^2) = f_{\frac{1}{2\sigma^2}, \frac{N}{2}}(\chi_N^2)} \quad (1.186)$$

On définit parfois

$$\chi^2 = \sum_{i=1}^N \frac{x_i^2}{\sigma^2} \quad (1.187)$$

Dans ce cas, un simple changement de variable donne

$$P_N(\chi^2) = \sigma^2 f_{\frac{1}{2\sigma^2}, \frac{N}{2}}(\chi^2) \quad (1.188)$$

Cette fonction est maximale pour $\chi^2 = N - 2$.

Retournons à notre question initiale qui était d'estimer la précision avec laquelle l'écart-type peut être obtenu d'un nombre limité de mesures. C'est la densité de probabilité pour la variable

$$s^2 = \sum_{i=1}^N \frac{x_i^2}{N} \quad (1.189)$$

qui nous intéresse alors. Encore une fois, en changeant de variable on trouve

$$P_N(s^2) = N f_{\frac{1}{2\sigma^2}, \frac{N}{2}}(s^2) \quad (1.190)$$

Cette fonction est tracée à la figure (1-13) où on a choisi $\sigma^2 = 1$ pour l'écart type de la densité de probabilité de la variable x_i . La courbe la plus étalée montre que pour $N = 4$ l'estimé le plus probable pour l'écart type est trop petit par rapport à la vraie réponse. En général, cette valeur la plus probable est $s^2 = \sigma^2 (1 - \frac{2}{N})$. À mesure que N augmente, l'estimé de l'écart type a une probabilité de plus en plus forte d'être correct. La courbe la plus piquée sur la figure est pour $N = 250$. La probabilité cumulative, illustrée à la figure (1-14) nous montre par exemple que pour $N = 4$, il y a environ neuf chance sur dix que notre estimé de l'écart type soit plus petit que deux fois le vrai écart type.

1.10 Résultats importants du chapitre

- La probabilité que deux événements complémentaires se produisent (OU) est la somme des probabilités de chacun des deux événements.
- La probabilité que deux événements statistiquement indépendants se produisent (ET) est le produit des probabilités de ces événements.
- Soit $P(u)$ la probabilité d'obtenir un résultat u pouvant prendre N valeurs discrètes. Cette distribution de probabilité doit être normalisée

$$\boxed{\sum_{u=1}^N P(u) = 1.} \quad (1.191)$$

La valeur moyenne de u est donnée par

$$\boxed{\bar{u} = \langle u \rangle = \sum_{u=1}^N P(u) u} \quad (1.192)$$

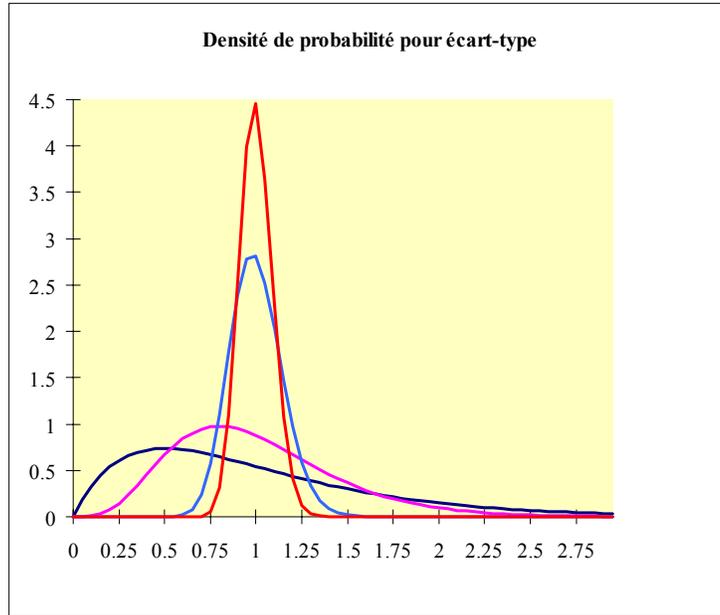


Figure 1-13 Densité de probabilité pour la valeur de $\sum_{i=1}^N x_i^2/N$ lorsque les x_i^2 sont obtenus d'une Gaussienne centrée à zéro dont l'écart type est égal à un. La courbe la moins piquée est pour $N = 4$ et les autres sont pour $N = 10$, $N = 100$, jusqu'à $N = 250$ pour la plus piquée.

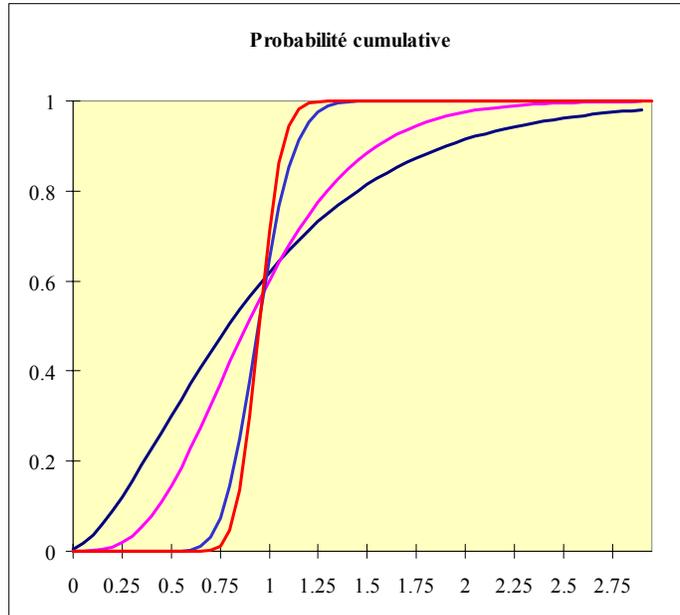


Figure 1-14 Probabilité pour que la valeur de $\sum_{i=1}^N x_i^2/N$ soit plus petite que la valeur indiquée sur l'abscisse lorsque les x_i^2 sont obtenus d'une Gaussienne centrée à zéro dont l'écart type est égal à un. La courbe la plus douce est pour $N = 4$ et les autres sont pour $N = 10$, $N = 100$, jusqu'à $N = 250$ pour la plus raide. (Le graphique a été obtenu par simple intégration par la règle du trapèze des résultats précédents plutôt qu'analytiquement),

On parle aussi de la moyenne de la distribution de probabilité $P(u)$. La variance de la même distribution est donnée par

$$\sigma^2 \equiv \langle (u - \langle u \rangle)^2 \rangle \quad (1.193)$$

alors que l'écart type est σ . Cette quantité est une mesure de la largeur de la distribution.

- La distribution binomiale

$$W_N(n_1) = \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} \quad (1.194)$$

décrit, par exemple, la probabilité de faire n_1 pas à droite dans une marche aléatoire de N pas. La valeur moyenne et la variance sont donnés, respectivement, par

$$\overline{n_1} = Np \quad (1.195)$$

$$\overline{(\Delta n_1)^2} \equiv \overline{n_1^2} - \overline{n_1}^2 = Npq$$

- Le truc utilisé pour obtenir la valeur moyenne et l'écart type de cette distribution, c'est-à-dire agir sur la condition de normalisation $(p+q)^N$ avec l'opérateur $(p\partial/\partial p)$, est un truc d'une très grande utilité. On applique une généralisation de ce truc dans le cas de la gaussienne et de façon plus générale en physique statistique lorsqu'on prend la dérivée de la fonction de partition pour calculer une valeur moyenne. (La fin de cette phrase ne deviendra compréhensible qu'au chapitre 6).

- Pour N grand, la distribution binomiale peut être approximée par

$$W_N(n_1) = \frac{1}{\sqrt{2\pi Npq}} \exp\left(-\frac{1}{2} \frac{(n_1 - Np)^2}{Npq}\right) \quad (1.196)$$

Cette forme fonctionnelle s'appelle une gaussienne. Ce résultat est obtenu en utilisant la formule de Stirling

$$\ln N! = N \ln N - N + \frac{1}{2} \ln(2\pi N) \quad (1.197)$$

et en développant $\ln W_N(n_1)$ autour de la valeur de n_1 pour laquelle $W_N(n_1)$ est maximum.

- C'est dans la limite du continu surtout qu'on parle d'une gaussienne. En général, $P(x)$ est une *densité* de probabilité gaussienne si elle a la forme

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}\right) \quad (1.198)$$

La probabilité que la variable aléatoire x soit comprise entre a et b se calcule ainsi

$$\int_a^b P(x) dx \quad (1.199)$$

La valeur moyenne est donnée par

$$\langle x \rangle = \int_{-\infty}^{\infty} xP(x) dx = \mu \quad (1.200)$$

et la variance par

$$\langle (x - \mu)^2 \rangle = \int_{-\infty}^{\infty} (x - \mu)^2 P(x) dx = \sigma^2. \quad (1.201)$$

- Pour un changement de variables général $\phi(u)$ où n valeurs de u , notées u_i ($i = 1, 2 \dots n$), correspondent à une valeur de ϕ donnée, le changement de variable correspondant pour la densité de probabilité est

$$\mathcal{W}(\phi) = \sum_{i=1}^n \mathcal{P}(u_i) \left| \frac{du}{d\phi} \right|_{u_i} \quad (1.202)$$

1.11 Problèmes pour le chapitre 1

1.11.1 Trouvez la bonne clé!²⁴

Un homme ne possédant pas toutes ses facultés essaie de rentrer chez lui. La porte est fermée à clé. Il possède un trousseau de six clés, dont une seule ouvre la porte. Il essaie une clé au hasard, puis remet le trousseau dans ses poches, le ressort et essaie encore une clé au hasard.

a) Quelle est la probabilité de ne pas encore avoir trouvé la bonne clé après avoir essayé N fois?

b) Quelle est la probabilité de ne pas avoir trouvé la bonne clé après $(N - 1)$ essais et de la trouver au N ième?

c) Si cet homme utilise la même brillante procédure à chaque soir, quel est le nombre moyen de fois qu'il essaie une clé avant de trouver la bonne? Vous aurez besoin du résultat suivant: $\sum_{n=0}^{\infty} p^n = (1 - p)^{-1}$. Ce résultat pour la somme d'une série géométrique est souvent utilisé et facile à obtenir en notant que si $S \equiv \sum_{n=0}^{\infty} p^n$ et la série converge, alors $pS = S - 1$.

1.11.2 Jeux de cartes

a) Faites une hypothèse sur la probabilité à priori de tirer une carte donnée d'un jeu de n cartes.

b) On prend deux cartes au hasard d'un jeu standard de 52 cartes.

- Quelle est la probabilité que la première soit un as de pic et la deuxième un as de coeur? Attention, la première carte n'est pas remise dans le tas avant de prendre la deuxième.
- Quelle est la probabilité de tirer un as de pic et un as de coeur mais dans n'importe quel ordre si la première carte est remise dans le tas avant de prendre la deuxième.

1.11.3 Le paradoxe des anniversaires!

Supposons que la probabilité à priori de fêter son anniversaire un jour donné soir $1/365$.

²⁴Reif. Prob. 1.5

a) Soient deux personnes prises au hasard. La probabilité que leur anniversaire ne soit pas la même date est $(1 - \frac{1}{365})$. Expliquez ce résultat à partir de la probabilité à priori donnée ci-haut.

b) Pour trois personnes, la probabilité qu'aucune n'ait son anniversaire le même jour est $(1 - \frac{1}{365}) \times (1 - \frac{2}{365})$. Pourquoi?

c) Montrez que pour une classe de 23 élèves, il y a plus d'une chance sur deux qu'au moins deux élèves fêtent leur anniversaire le même jour.

1.11.4 Des juges impartiaux?

Six juges doivent classer vingt candidats. Chacun d'entre eux connaît bien un des candidats. Une fois la compétition terminée, on s'aperçoit que les six candidats connus des juges se retrouvent dans les dix premières places. Quelle est la probabilité que ceci se soit produit par hasard et que les juges aient vraiment été impartiaux?

1.11.5 Le triangle de Pascal, d'intérêt historique

On prend l'équation suivante comme définition des nombres C_n^N

$$(p + q)^N = \sum_{n=0}^N C_n^N p^n q^{N-n}.$$

a) En utilisant le fait que $(p + q)^N (p + q) = (p + q)^{N+1}$ trouvez une relation entre les C_n^N et les $C_{n'}^{N-1}$ et montrez qu'on peut résumer le résultat à l'aide de l'algorithme suivant, dit "triangle de Pascal".

$$\begin{array}{cccccc}
 & & & & & 1 \\
 & & & & 1 & & 1 \\
 & & & 1 & 2 & 1 & \\
 & & 1 & 3 & 3 & 1 & \\
 & 1 & 4 & 6 & 4 & 1 & \\
 1 & 5 & 10 & 10 & 5 & 1 &
 \end{array}$$

Algorithme: Chaque nombre est la somme des deux nombres immédiatement au-dessus de lui à gauche et à droite. Le nombre C_n^N apparaît dans la n ième position de la $(N + 1)$ ième rangée.

b) Montrez que $C_n^N = \frac{N!}{n!(N-n)!}$ satisfait l'algorithme du triangle de Pascal.

1.11.6 Mouvement Brownien

Une particule se déplace en une dimension d'une distance ℓ pendant un temps τ entre chaque collision. À chacune de ces collisions, elle change de direction, allant à droite avec une probabilité $p = 1/2$ ou à gauche avec une probabilité $q = 1/2$. Utilisant ce que vous savez sur la binomiale, (valeur moyenne, variance...) démontrez que

$$\langle x \rangle = 0 \tag{1.203}$$

$$\langle x^2 \rangle = Dt \quad (1.204)$$

où x est la distance parcourue depuis le point de départ à $x = 0$ au temps $t = 0$ et le nombre de collisions fait au temps t est $N = t/\tau$. La constante de diffusion D est donnée par

$$D = \ell^2/\tau \quad (1.205)$$

1.11.7 Moyenne et probabilités dans un jeu de hasard

Un de vos amis vous propose de jouer deux cents fois avec lui au jeu suivant. Vous lancez deux dés. Si vous faites un sept, il vous donne 4.00\$ et si vous ne faites pas un sept, vous lui donnez 1.00\$. Combien croyez-vous gagner ou perdre après deux cents lancers. Accepterez vous de jouer? Votre ami l'est-il toujours?

1.11.8 Erreur sur la moyenne

Soit la probabilité $P(i)$ pour une variable aléatoire discrète i . Par exemple, $P(i)$ pourrait être la probabilité de rouler la valeur i dans le jeu de dés précédent. Soit le gain $g(i)$ pour une valeur i des dés. Le gain moyen est donné par

$$\bar{g} = \sum_{i=2}^{12} g(i) P(i) \quad (1.206)$$

et l'écart quadratique moyen par

$$\sigma^2 = \overline{g^2} - \bar{g}^2 \quad (1.207)$$

Maintenant, considérons une autre variable aléatoire, soit le gain après N parties, normalisé par le nombre de parties.

$$G = \frac{1}{N} \sum_{n=1}^N g_n \quad (1.208)$$

- Quelle est la distribution de probabilité pour cette variable aléatoire? (Produit des probabilités)
- Obtenez la valeur moyenne de G en fonction de \bar{g} .
- Montrez que l'écart type de la distribution de G est plus petite que l'écart type de la distribution de g par un facteur $1/\sqrt{N}$.
- Quelle est l'erreur sur l'estimé de vos gains ou pertes dans le problème précédent avec votre ami? Cela renforce-t-il votre décision?

1.11.9 Calculs avec la gaussienne

Un expérimentateur donne un tableau de différentes mesures expérimentales de la même quantité x . Il vous dit que la valeur moyenne de x est ℓ et que l'écart type

de l'ensemble des mesures est σ . Ceci veut dire que la probabilité qu'une mesure donnée se retrouve entre la valeur x et la valeur $x + dx$ est donnée par

$$P(x) dx = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-(x - \ell)^2 / (2\sigma^2)\right] dx$$

a) Montrez que cette distribution de probabilité est normalisée.

b) Vérifiez la valeur moyenne $\langle x \rangle$ et l'écart type $\sqrt{\langle (x - \langle x \rangle)^2 \rangle}$ de cette distribution de probabilité.

c) En regardant les mesures expérimentales, vous notez qu'environ une mesure sur trois se situe en dehors de l'écart type. Est-ce cohérent avec l'hypothèse de départ?

d) Calculer la probabilité que la mesure soit comprise *i*) entre $-\infty$ et $\ell - \sigma$, *ii*) entre $\ell - \sigma$ et $\ell + \sigma$ et *iii*) entre $\ell + \sigma$ et ∞ .

1.11.10 Quelques statistiques de radioactivité²⁵

Des particules α sont émises par une source radioactive pendant un intervalle de temps T . Cet intervalle de temps T est beaucoup plus petit que le temps de vie de la substance radioactive, de telle sorte qu'on pourra supposer que le nombre d'atomes qui ne sont *pas* encore désintégrés durant l'intervalle T demeure très bien approximé par une constante. Considérons maintenant un sous-intervalle de temps $\Delta t \ll T$. Puisque les particules sont émises au hasard, le nombre de désintégrations se produisant durant un intervalle Δt est indépendant des désintégrations qui se sont produites à d'autres temps. De plus, on peut choisir Δt tellement petit que la probabilité de plus d'une désintégration dans le temps Δt est négligeable. On a donc une probabilité $p \ll 1$ d'avoir une désintégration dans l'intervalle Δt . On a cependant $N = T/\Delta t$ "essais" durant notre période d'observation T . En répétant l'expérience un grand nombre de fois pour la période T on sait que le nombre moyen de désintégrations durant cet intervalle est donné par $Np = \lambda$. Le nombre λ est un nombre fini, relativement petit.

a) Expliquez pourquoi la distribution de probabilité pour que n désintégrations se produisent durant l'intervalle T est une binomiale

$$\frac{N!}{n!(N-n)!} p^n q^{N-n} \quad (1.209)$$

On montrera, en suivant les étapes ci-dessous, que dans la limite $p \rightarrow 0$, $N \rightarrow \infty$, $Np = \lambda$, cette distribution prend la forme dite de Poisson,²⁶

$$W(n, \lambda) = \frac{\lambda^n}{n!} e^{-\lambda} \quad (1.210)$$

b) Comme N est grand alors que n/N et p sont petits, justifiez la séquence d'approximations suivantes à l'ordre dominant en p et en n/N

$$q^{N-n} = (1-p)^{N-n} = \exp[(N-n) \ln(1-p)] \approx \exp[p(n-N)] \quad (1.211)$$

$$\approx \exp(-Np) = \exp(-\lambda) \quad (1.212)$$

²⁵Reif, Prob. 1.12 et 1.9

²⁶Reif, Prob. 1.9

Remarque 24 On retrouve continuellement ce genre d'approche en théorie des probabilités. Elle est de loin préférable au développement en série direct. En effet, supposons qu'on veuille trouver la limite N grand de

$$(1 - p)^N \quad (1.213)$$

lorsque p est petit. Il n'est pas très fructueux de faire le développement direct

$$(1 - p)^N \simeq 1 - Np + \dots \quad (1.214)$$

car $Np \gg 1$ veut dire qu'on a dépassé le rayon de convergence de la série.

c) Comme N est grand, on peut utiliser la formule de Stirling pour les factorielles où N apparaît. Plus spécifiquement, on a

$$N! \approx \exp [N \ln N - N] \quad (1.215)$$

$$(N - n)! \approx \exp [(N - n) \ln (N - n) - N + n] \quad (1.216)$$

Montrez à partir de cette approximation de Stirling qu'à l'ordre dominant en n/N ,

$$\frac{N!}{n!(N - n)!} \approx \frac{N^n}{n!} \quad (1.217)$$

d) Utilisant les résultats de b) et c), montrez que dans cette limite, la binomiale devient la distribution de Poisson.

e) Montrez que la distribution de Poisson est normalisée, c'est-à-dire que

$$\sum_{n=0}^{\infty} W(n, \lambda) = 1 \quad (1.218)$$

f) Prouvez que

$$e^{-\lambda} \lambda \frac{\partial}{\partial \lambda} [W(n, \lambda) e^{\lambda}] = nW(n, \lambda) \quad (1.219)$$

g) Utilisez les résultats de e) et de f) pour montrer que la valeur moyenne de la distribution de Poisson est donnée par,

$$\langle n \rangle = \sum_{n=0}^{\infty} nW(n, \lambda) = \lambda \quad (1.220)$$

h) Par des moyens semblables, calculez l'écart quadratique moyen $\langle n^2 \rangle - \langle n \rangle^2$ pour la distribution de Poisson.

1.11.11 Les erreurs de frappe²⁷

Supposons qu'on trouve 625 erreurs de typographie au total dans un livre de 500 pages. Quelle est la probabilité d'avoir deux erreurs dans une page? Expliquez votre choix de distribution.

²⁷Reif, Prob. 1.11

1.11.12 Élimination du bruit²⁸

Pour observer la surface d'une planète, il est très utile d'envoyer des signaux radar pendant un temps disons τ , et de détecter ce signal plus tard pendant le même temps τ . Le signal a qui nous revient, est souvent corrompu par du bruit, en d'autres mots

$$a = a_s + a_b$$

où a_s est le signal qu'on veut vraiment détecter, et a_b est le bruit. Supposons qu'on soit dans un cas où le bruit est beaucoup plus fort que le signal, c'est-à-dire

$$\sqrt{\langle (\Delta a_b)^2 \rangle} \gg a_s.$$

La situation n'est pas désespérée quand même. Montrez que si $\langle a_b \rangle = 0$, alors en envoyant et détectant plusieurs fois le même signal radar, on peut finalement extraire le signal du bruit. Estimez le nombre de fois qu'il faudra envoyer le signal radar sachant l'intensité du bruit et du signal à détecter.

1.11.13 Changements de variables²⁹

Soit un ensemble d'oscillateurs harmoniques de masse m et de constante de ressort k . On sait que l'énergie de chacun des oscillateurs est la même, soit E , mais que le temps auquel chacun a été démarré est aléatoire (mais distribué uniformément, c'est-à-dire sans qu'un temps soit plus probable que l'autre). Trouvez la densité de probabilité $p(x)$ telle que $p(x) dx$ est la probabilité que la position d'une des masses soit dans l'intervalle entre x et $x + dx$.

1.11.14 Mesures expérimentales

La Fig.(1-15) montre les résultats d'une mesure primitive de distance.

- Quelle est la valeur moyenne mesurée?
- Quel est l'écart type de la distribution des mesures?
- Estimez l'erreur sur la valeur de la moyenne.

1.11.15 Configurations de bosons

a) Soit un ensemble de n barres verticales identiques et de p objets identiques. Montrez qu'il y a

$$\frac{(p+n)!}{n!p!}$$

façons de les ordonner si on ne compte pas les permutations des objets entre eux et les permutations des barres verticales entre elles. Pour être plus spécifique, les deux rangées ci-dessous sont deux façons différentes d'ordonner les objets et les

²⁸Reif, Prob. 1.21

²⁹Q2001

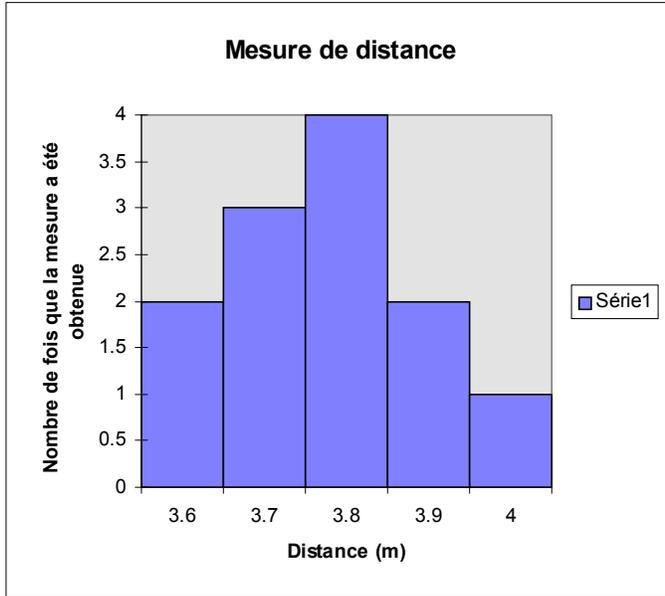


Figure 1-15 Résultats de mesures primitives de distance.

barres verticales, mais si deux barres verticales ($| \quad |$) sont permutées entre elles, ou si deux objets ($o \quad o$) sont permutés entre eux dans une rangée, l'ordre n'est pas considéré comme différent.

$$\begin{array}{cccccccc}
 o & | & o & o & | & o & o & o & \dots \\
 | & | & o & | & o & o & | & o & \dots \\
 \vdots & \vdots
 \end{array}$$

b) Soit un ensemble de N particules identiques. Utilisant le résultat précédent, combien y a-t-il de façons de répartir ces N particules dans m boîtes (chacune des “boîtes” en pratique est un niveau d'énergie quantique). Chaque boîte peut contenir un nombre arbitraire de particules. De telles particules identiques s'appellent bosons. Toutes les particules élémentaires (incluant le photon) sont soit des bosons, soit des fermions!

1.11.16 Configurations de fermions

Les fermions sont aussi des particules identiques. Cependant, contrairement aux bosons, on ne peut pas en mettre plus d'une par niveau d'énergie (“boîte”). Si on a M niveaux d'énergie et N fermions, avec $M > N$, combien y a-t-il de façons de répartir les N particules dans les M niveaux d'énergie? Le résultat est différent du cas des boson. Un niveau d'énergie est soit vide, soit occupé.

1.11.17 Erreur quadratique moyenne

Montrez que pour toute distribution de probabilité,

$$\langle (u - \langle u \rangle)^2 \rangle = \langle u^2 \rangle - \langle u \rangle^2 \quad (1.221)$$

1.11.18 Pile ou face truqué

Supposons qu'une pièce de monnaie truquée tombe du côté pile avec une probabilité $p = 0.55$. Si vous gagnez 1\$ à chaque fois que la pièce tombe sur pile et que vous perdez 1.02\$ à chaque fois qu'elle tombe sur face,

- a) Quel est votre gain moyen dans une partie composée de 10 lancers?
- b) Quel est votre gain moyen dans une partie composée de 100 lancers?
- c) Si vous ne pouvez faire qu'une partie, choisirez vous une partie de 10 lancers ou de 100 lancers et pourquoi? Quantifiez votre réponse en discutant de l'écart type.

1.11.19 Sondages

Vous désirez faire un sondage pour savoir l'opinion de l'électorat sur une question référendaire hypothétique. Combien de personnes devrez-vous interroger pour obtenir avec une précision de l'ordre de $\pm 3\%$ le pourcentage de personnes qui voteront pour le OUI ou pour le NON. Vous pouvez supposer que le vote sera très serré.