

Nom :

Examen de Statistiques

Septembre 2004

1 Premier problème

La série temporelle donnée en Table (1) représente la hauteur des marées (en mètres) à Dunkerke mesurée chaque deux heures pendant la journée du 31 août 2004 (source <http://maree.frbateaux.net/3>).

00H00	3.3	08H00	1.3	16H00	5.6	(1)
02H00	6.1	10H00	1	18H00	3.5	
04H00	5.5	12H00	2.8	20H00	1.3	
06H00	3.2	14H00	5.9	22H00	0.4	

1. Donner les définitions et calculer : la moyenne et la variance.
2. Y a-t-il des corrélations évidentes dans cette série temporelle ?
3. Peut-on prendre ces données comme étant obtenus avec un générateur de nombres aléatoires ?

2 Deuxième problème

Pour la distribution de probabilité présentée dans la Fig. 1:

1. Marquer sur la figure la moyenne, la variance et le mode de $p(x)$.
2. A votre avis, quelle est la forme fonctionnelle des picks ?

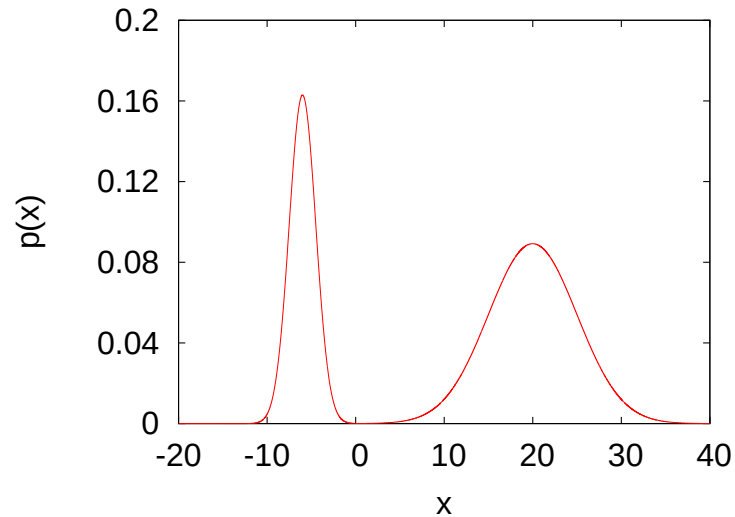


Figure 1: La distribution de probabilité $p(x)$ du problème 2.

3 Troisième problème

Prenons un jeu de 52 cartes sans joker. Leurs couleurs sont pique, coeur, carreau et trèfle.

1. Quelle est la probabilité de tirer un coeur ?
2. Quelle est la probabilité de tirer un roi ?
3. Quelle est la probabilité de tirer le roi de coeur ?
4. Quelle est la probabilité de tirer un roi ou un coeur ? Trouver le résultat en utilisant deux arguments différents.
5. Les événements ‘tirer un roi’ et ‘tirer un coeur’ sont-ils indépendants ?
6. Répondre la dernière question si l’on rajoute un joker au jeu de cartes. Expliquer le résultat trouvé.

4 Quatrième problème

Après effectuer une mesure on trouve les résultats en Table. 4 pour la variable indépendante x et la variable dépendante y .

x	y
0.05	0.0
0.95	3.5
2.10	18
3.05	37
3.95	63

1. Trouver l'expression théorique de la meilleur regression linéaire.
2. Calculer les coefficients inconnus, qu'on appellera a et b , pour les données en Table 1.
3. Avez-vous trouvé un résultat de bonne qualité ?
4. Quelle solution pouvez vous proposer ?

5 Cinquième problème

On calculera l'intervale de confiance pour la proportion c d'une population qui votera le candidat C dans une election.

Après avoir sondé une population de taille n on trouve que x membres de l'échantillon voteront pour C .

1. Quelle est la proportion y de l'échantillon que votera pour C ?
2. Est y une variable déterministe ou aléatoire ?
3. Quand le nombre n est très grand on supposera que y est donnée par une distribution normale. Quelle sont la moyenne et la variance de y ?

4. On cherche à déterminer c avec 95% de confiance.

Donner la formule analytique pour l'intervale des valeurs de c déterminés avec cette confiance.

Sachant que pour une distribution normale $a = 1.96$ si $p(|z| < a) = 0.95$, donner la valeur numérique de l'intervale de confiance pour $n = 200$ et $X = 124$.

6 Solution

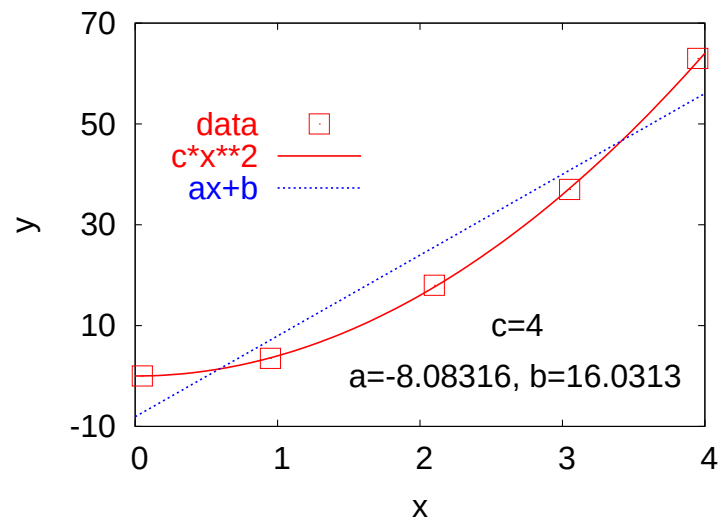


Figure 2: Description graphique du problème 4.