

Examen Final : MST 2

19 mai 2017

Durée : 1h45 ; Aucun document autorisé ; pas de calculatrice.

La qualité de la rédaction est prise compte. Lire l'intégralité de l'énoncé avant de démarrer.

Attention : Enoncé Recto-Verso

Exercice 1:

Soit $(X_m)_{m \geq 1}$ une suite de variables aléatoires de Bernoulli de même loi $B(1, p)$ ($p \in]0, 1[$) et indépendantes, et soit $r \geq 1$ un entier. On note Y la variable aléatoire correspondant au nombre de tirages consécutifs nécessaires X_m pour que le nombre d'échecs (de "0") soit égale à r . La loi de Y est appelée la loi *Binomiale Négative* (notée $BN(p, r)$), de paramètres p et r .

1. Démontrer que pour tout $k \geq 1$,

$$P(Y = k) = p^k (1 - p)^r \binom{k + r - 1}{k}$$

où $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ est le coefficient binomial.

2. Dans un premier temps, on suppose que le paramètre r est connu. Est ce que le modèle de loi binomiale est une famille exponentielle ? En déduire que $E_{(p,r)}[Y] = \frac{pr}{1-p}$ et que la variance $V_{(p,r)}(Y) = \frac{pr}{(1-p)^2}$.
3. Soit $y_1, y_2, \dots, y_n$ un échantillon iid de loi commune $BN(p, r)$. Ecrire la log-vraisemblance $\mathcal{L}(y_1, \dots, y_n; p, r)$ de cette échantillon. Il pourra être pratique d'utiliser le fait que le coefficient binomial $\binom{k+r-1}{k}$ est aussi égale à $\frac{\Gamma(k+r)}{k!\Gamma(r)}$, où Γ est la fonction définie par $\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt$ (en effet, $\Gamma(k) = (k-1)!$, pour tout entier $k \geq 1$).
4. En déduire l'estimateur du maximum de vraisemblance $\hat{p}$ de p .
5. Calculer l'information de Fisher $I(p)$ du modèle. Que représente $I(\hat{p})$?
6. Pour $n = 100$, construire un intervalle de confiance de niveau 95% approximativement, en justifiant votre approximation.
7. On suppose maintenant que le paramètre r est lui aussi inconnu. On suppose de plus que le coefficient r est réel positif (et n'est plus seulement un entier). Ceci est possible, car le coefficient binomial $\binom{k+r-1}{k}$ se généralise aux réels grâce à la fonction Gamma.
Comment estimez vous simultanément les 2 paramètres (p, r) ? Comment faites vous pour résoudre ce problème pour lequel il n'y a pas de solution explicite ?
8. Calculer la matrice de variance (asymptotique) de l'estimateur du maximum de vraisemblance $(\hat{p}, \hat{r})$.
9. Construire une ellipse de confiance pour les 2 paramètres de niveau $1 - \alpha$ pour $n = 100$?

P 1/2 Tournez la page SVP

Exercice 2

Soit $X_1, \dots, X_n$ un échantillon i.i.d dont la loi appartient au modèle gaussien $\{N(0, \sigma^2), \sigma > 0\}$, et soit $\sigma_0 < \sigma_1$. On veut construire le test statistique

$$\begin{cases} H_0 : \sigma = \sigma_0 \\ H_1 : \sigma = \sigma_1 \end{cases}$$

de risque de 1ère espèce α .

1. Rappeler l'approche de Neyman-Pearson pour la construction d'un test simple. Qu'est ce que cela garantit ? On définira rigoureusement toutes les quantités impliquées.
2. Construire le test de niveau α .
3. Calculer la puissance.
4. On a un échantillon de taille $n = 10$ tel que $\sum_{i=1}^{10} x_i = 3$, et $\sum_{i=1}^{10} x_i^2 = 11$. On teste l'hypothèse $\sigma = 0.1$ contre $\sigma = 0.3$, avec un risque $\alpha = 5\%$: rejetez vous l'hypothèse H_0 ? Ci-dessous le tableau de quelques quantiles q_α de la distribution du χ_n^2 à $n = 10$ et $n = 9$ degrés de libertés.

α	2.5%	5%	95%	97.5%
$P(\chi_{10}^2 \leq q_\alpha) = \alpha$	3.24	3.94	18.3	20.48
$P(\chi_9^2 \leq q_\alpha) = \alpha$	2.70	3.32	16.91	19.02

5. On suppose maintenant que les observations s'écrivent $X_i = \sigma Y_i$ avec Y_i suivant loi normale $N(1, 1)$ (et Y_i échantillon i.i.d). Ecrire la vraisemblance de l'échantillon pour ce nouveau modèle.
6. On souhaite tester dans ce nouveau modèle toujours la même hypothèse

$$\begin{cases} H_0 : \sigma = \sigma_0 \\ H_1 : \sigma = \sigma_1 \end{cases} \quad \sigma_0 < \sigma_1$$

Déterminer la zone de rejet et la statistique de test dans ce cas. A votre avis, comment faire pour conclure ?

Exercice 3

Soient $(X_1, X_2, \dots, X_n)$ et $(Y_1, Y_2, \dots, Y_m)$ 2 échantillons indépendants, de loi de Bernoulli $B(1, p)$ et $B(1, q)$ respectivement ($P(X_i = 1) = p$, et $P(Y_i = 1) = q$), avec de paramètre $0 < p, q < 1$.

1. Trouver les estimateurs du maximum de vraisemblance pour p et q
2. Dédire des estimateur des variances des estimateurs et en déduire un intervalle de confiance (approchée) à 90% pour p et q respectivement.
3. Proposer un estimateur pour la différence $p - q$ et proposer un intervalle de confiance (approchée) à 90%.