Sorbonne Université | 9 décembre 2024
Master Mathématiques et Applications | Aucun document autorisé
MU4MA015 : Statistique | Durée : 1 heure

Contrôle Continu

Rappels

- Il sera tenu compte du soin apporté à la présentation et à la rédaction.
- Si vous n'arrivez pas à démontrer un résultat, vous pouvez l'admettre pour la suite de l'exercice.
- Si $T \sim \Gamma(a,b)$ avec $a > 2$ et $b > 0$, alors $\mathbb{E}[1/T] = \frac{b}{a-1}$ et $\mathbb{E}[1/T^2] = \frac{b^2}{(a-1)(a-2)}$.

Exercice 1 (6 points)

Soit $n \geq 3$. On considère le cadre bayésien

$$\boldsymbol{\theta} \sim \Pi = \mathcal{E}(1)$$
$$\mathbf{X}|\boldsymbol{\theta} \sim P_{\boldsymbol{\theta}}^{\otimes n} = \mathcal{E}(\boldsymbol{\theta})^{\otimes n}.$$

1. Donner la loi a posteriori.
Correction. On a
$$\pi(\theta|\mathbf{X}) \propto e^{-\theta}\mathbf{1}_{\theta>0}\prod_{i=1}^{n}\theta e^{-\theta X_i} \propto \theta^n e^{-(n\overline{X}_n+1)\theta}\mathbf{1}_{\theta>0}$$
donc $\boldsymbol{\theta}|\mathbf{X} \sim \Gamma\left(n+1, n\overline{X}_n+1\right)$.
2. On considère la perte
$$\ell(\theta,t) = e^{\theta}(t-\theta)^2.$$
Donner un estimateur de Bayes $T^\star = T^\star(\mathbf{X})$ pour Π et cette fonction de perte.
Correction. On a
$$\rho(\Pi, T|\mathbf{X}) \propto \int_{\mathbb{R}} (T-\theta)^2 \underbrace{\theta^n e^{-\theta n\overline{X}_n}\mathbf{1}_{\theta>0}}_{\propto f_{\Gamma(n+1,n\overline{X}_n)}(\theta)} d\theta$$
donc
$$T^* = \mathbb{E}[\Gamma(n+1, n\overline{X}_n)|\mathbf{X}] = \frac{n+1}{n\overline{X}_n}$$
est un estimateur de Bayes.
3. Montrer que, sous $P_\theta^{\otimes n}$, la variable $n\overline{X}_n$ suit une loi Gamma dont on précisera les paramètres. Grâce au rappel, en déduire $\mathbb{E}_\theta\left[\frac{1}{n\overline{X}_n}\right]$ et $\mathbb{E}_\theta\left[\frac{1}{(n\overline{X}_n)^2}\right]$.
Correction. Sous $P_\theta^{\otimes n}$, c'est-à-dire sachant $\boldsymbol{\theta} = \theta$, les X_i sont i.i.d. exponentielles de paramètre θ donc $n\overline{X}_n \sim \Gamma(n,\theta)$ et par conséquent
$$\mathbb{E}_\theta\left[\frac{1}{n\overline{X}_n}\right] = \frac{\theta}{n-1} \qquad \text{et} \qquad \mathbb{E}_\theta\left[\frac{1}{(n\overline{X}_n)^2}\right] = \frac{\theta^2}{(n-1)(n-2)}.$$

4. Montrer que le risque ponctuel $R(\theta,T^\star)=\mathbb{E}_\theta[\ell(\theta,T^\star)]$ peut se mettre sous la forme

$$R(\theta,T^\star)=\frac{an+b}{(n-1)(n-2)}\theta^2e^\theta,$$

où a et b sont deux constantes positives que l'on précisera.
Correction. On a

$$R(\theta,T^\star)=\mathbb{E}_\theta\left[e^\theta(T^\star-\theta)^2\right]=e^\theta\mathbb{E}_\theta\left[(T^\star-\theta)^2\right]=e^\theta\mathbb{E}_\theta\left[\left(\frac{n+1}{n\overline{X}_n}-\theta\right)^2\right],$$

d'où

$$R(\theta,T^\star)=e^\theta\left((n+1)^2\mathbb{E}_\theta\left[\frac{1}{n^2\overline{X}_n^2}\right]-2(n+1)\mathbb{E}_\theta\left[\frac{1}{n\overline{X}_n}\right]\theta+\theta^2\right),$$

et, après simplification,

$$R(\theta,T^\star)=\frac{n+7}{(n-1)(n-2)}\theta^2e^\theta,$$

5. Que vaut le risque de Bayes $R_B(\Pi)$? En déduire le risque minimax R_M.
Correction. On a

$$R_B(\Pi)=\mathbb{E}\left[R(\boldsymbol{\theta},T^\star)\right]=\int_{\mathbb{R}}R(\theta,T^\star)\pi(\theta)d\theta=\int_0^{+\infty}\frac{n+7}{(n-1)(n-2)}\theta^2e^\theta e^{-\theta}d\theta=+\infty.$$

Puisque le risque de Bayes est plus petit que le risque minimax, on en déduit que $R_M=+\infty$.
6. Pouvez-vous proposer un autre estimateur de Bayes ?
Correction. Tout estimateur, i.e. toute fonction $T(\mathbf{X})$, est donc de Bayes.
7. On souhaite tester

$$H_0:\boldsymbol{\theta}=1\quad\text{contre}\quad H_1:\boldsymbol{\theta}=2\,.$$

Pour cela, on considère maintenant la loi a priori $\Pi=\frac{1}{2}\delta_1+\frac{1}{2}\delta_2$ et la fonction de perte équilibrée. Montrer que le test de Bayes peut se mettre sous la forme $\varphi^*(\mathbf{X})=\mathbb{1}_{\{\overline{X}_n\leq c\}}$, pour une certaine constante $c\in\mathbb{R}$ que l'on déterminera.
Correction. Par le cours, on sait qu'un test de Bayes pour la perte équilibrée est donné par

$$\varphi^*(\mathbf{X})=\mathbb{1}_{\{\Pi(\{1\}|\mathbf{X})\leq\frac{1}{2}\}},$$

or

$$\Pi(\{1\}|\mathbf{X})=\frac{\Pi(\{1\})p_1(\mathbf{X})}{\Pi(\{1\})p_1(\mathbf{X})+\Pi(\{2\})p_2(\mathbf{X})}=\frac{\frac{1}{2}e^{-n\overline{X}_n}}{\frac{1}{2}e^{-n\overline{X}_n}+\frac{1}{2}2^ne^{-2n\overline{X}_n}}$$

d'où

$$\varphi^*(\mathbf{X})=\mathbb{1}_{\{e^{-n\overline{X}_n}\leq 2^ne^{-2n\overline{X}_n}\}}=\mathbb{1}_{\{\overline{X}_n\leq\ln(2)\}}\,.$$

Exercice 2 (4 points)

On considère, dans cet exercice, un modèle linéaire gaussien

$$Y=X\beta+\varepsilon=\begin{pmatrix}1 & t_1\\ \vdots & \vdots\\ 1 & t_n\end{pmatrix}\beta+\varepsilon,$$

où, comme d'habitude,

— Y est un vecteur aléatoire de n observations $Y_1, \dots, Y_n$;
— X est une matrice de mesures connues et déterministes, de taille $n \times 2$ et de rang 2 ;
— β est un vecteur déterministe inconnu de taille 2 ;
— ε est un vecteur aléatoire d'erreurs inconnues avec $\varepsilon \sim \mathcal{N}(0, \sigma^2 I_n)$ où $\sigma^2 > 0$ est inconnu.

1. Par quoi faut-il remplacer `???` pour que le script suivant construise une réalisation `y` du vecteur aléatoire Y.

[3]:
```
n = 100
noise = stats.norm(scale=true_sigma2())
t = measure_t(n)
beta = true_beta()
X = np.column_stack((np.ones(n), t))
y = ???
```

Correction. `X @ beta + noise.rvs(size=n)`

2. Étant donné les informations suivantes sur la matrice X et sur la réalisation $y = Y(\omega)$, calculer l'estimateur des moindres carrés de β, noté $\widehat{\beta}$, et donner sa réalisation $\widehat{\beta}(\omega)$.

$$\sum_{i=1}^{n} t_i = 0, \qquad \sum_{i=1}^{n} t_i^2 = 400, \qquad \sum_{i=1}^{n} t_i y_i = 100, \qquad \sum_{i=1}^{n} y_i = 100, \qquad \widehat{\sigma}^2 = 1.$$

Correction. La sortie précédente donne $n = 100$. De plus, en notant $t \in \mathbb{R}^n$ le vecteur des mesures `t`, on a $X = [\mathbb{1}|t] \in \mathbb{R}^{n\times 2}$, donc $X^\top X = \begin{pmatrix} n & 0 \\ 0 & \|t\|^2 \end{pmatrix} = \begin{pmatrix} 100 & 0 \\ 0 & 400 \end{pmatrix}$ et $X^\top Y = \begin{pmatrix} \mathbb{1}^\top Y \\ t^\top Y \end{pmatrix} = \begin{pmatrix} 100 \\ 100 \end{pmatrix}$. Puisque $\widehat{\beta} = (X^\top X)^{-1} X^\top Y$, on en conclut que $\widehat{\beta}(\omega) = \begin{pmatrix} 1 \\ 1/4 \end{pmatrix}$.

3. Que produit le script suivant (justifier par des éléments théoriques) ? Le résultat peut-il être lu dans le résumé des résultats donné ensuite via la commande `res.summary()` ?

[10]:
```
model = sm.OLS(y, X)
res = model.fit()
```

[11]:
```
level = .05
stu = stats.t(df=n-2)
ic_bound = np.sqrt(res.scale / np.diag(X.T @ X)) * stu.ppf(1 - level/2)
for c, s in zip(res.params, ic_bound):
    print(f"[{c-s:0.3f}, {c+s:0.3f}]")
```

```
[0.802, 1.198]
[0.151, 0.349]
```

[12]:
```
res.summary()
```

[12]:

Dep. Variable :	y	**R-squared :**	0.203
Model :	OLS	**Adj. R-squared :**	0.195
Method :	Least Squares	**F-statistic :**	25.00
Date :	Fri, 29 Nov 2024	**Prob (F-statistic) :**	2.51e-06
Time :	11 :20 :50	**Log-Likelihood :**	-140.88
No. Observations :	100	**AIC :**	285.8
Df Residuals :	98	**BIC :**	291.0
Df Model :	1		
Covariance Type :	nonrobust		

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
const	1.0000	0.100	10.000	0.000	0.802	1.198
x1	0.2500	0.050	5.000	0.000	0.151	0.349

Omnibus :	4.728	**Durbin-Watson :**	2.206
Prob(Omnibus) :	0.094	**Jarque-Bera (JB) :**	5.178
Skew :	-0.239	**Prob(JB) :**	0.0751
Kurtosis :	4.007	**Cond. No.**	2.00

Correction. Puisque `res.params` = $\widehat{\beta}(\omega)$, `res.scale` = $\widehat{\sigma}^2(\omega)$ et `1 / np.diag(X.T @ X)` = $([(X^\top X)^{-1}]_{11}, [(X^\top X)^{-1}]_{22})$, il devient clair que `ic_bound` est le vecteur des rayons des réalisations des intervalles de confiance de β_1 et β_2, qui sont affichés ensuite et dont on rappelle la forme pour β_1 :

$$\left[\widehat{\beta}_1(\omega) \pm \sqrt{\widehat{\sigma}^2(\omega)[(X^\top X)^{-1}]_{11}} F^{-1}_{\mathcal{T}_{n-2}}(1-\alpha/2)\right],$$

pour $\alpha \in]0,1[$ et $F_{\mathcal{T}_{n-2}}$ la fonction de répartition de $\mathcal{T}_{n-2}$.

Oui, les réalisations de ces deux intervalles de confiance peuveut être lues dans la colonne `[0.025    0.975]`.

4. On rappelle que l'ellipse centrée en $(x_0, y_0) \in \mathbb{R}^2$ et définie par :

$$\left\{(x', y') \in \mathbb{R}^2, \frac{(x'-x_0)^2}{a^2} + \frac{(y'-y_0)^2}{b^2} \leq 1\right\},$$

(où $a > 0$ et $b > 0$) a pour aire πab.

Commenter le script suivant (utiliser des éléments théoriques) et son résultat.

[13]:
```
fish = stats.f(dfn=2, dfd=n-2)

box_bound = np.sqrt(res.scale / np.diag(X.T @ X)) * stu.ppf(1 - level/4)

print(f"Aire 1 : {np.prod(2 * box_bound):.3f}")
print(f"Aire 2 : {np.pi / np.prod( np.sqrt(np.diag(X.T @ X) / (2 * res.scale * fish.ppf(1 - level))) ):.3f}")
```

```
Aire 1 : 0.104
Aire 2 : 0.097
```

Correction. D'après ce qui précède, `Aire 1` correspond à l'aire de la réalisation du rectangle de confiance de niveau $1-\alpha$ pour β : on a effectivement pris les quantiles d'ordre $1-\alpha/4$ en lien avec l'application de la borne de l'union vue en cours (également appelée méthode de Bonferroni).

`Aire 2` correspond à l'aire de la réalisation de l'ellipse de confiance de niveau $1-\alpha$:

$$\left\{\beta' \in \mathbb{R}^2, (\widehat{\beta}(\omega) - \beta')^\top \frac{X^\top X}{2\widehat{\sigma}^2(\omega) F^{-1}_{\mathcal{F}^2_{n-2}}(1-\alpha)} (\widehat{\beta}(\omega) - \beta') \leq 1\right\},$$

où $F_{\mathcal{F}^2_{n-2}}$ est la fonction de répartition de $\mathcal{F}^2_{n-2}$ et où l'on a remarqué que, pour tout $\beta' \in \mathbb{R}^2$:

$$\begin{aligned}&(\widehat{\beta}(\omega) - \beta')^\top \frac{X^\top X}{2\widehat{\sigma}^2(\omega) F^{-1}_{\mathcal{F}^2_{n-2}}(1-\alpha)} (\widehat{\beta}(\omega) - \beta') \\ &= (\widehat{\beta}_1(\omega) - \beta'_1)^2 \frac{[X^\top X]_{11}}{2\widehat{\sigma}^2(\omega) F^{-1}_{\mathcal{F}^2_{n-2}}(1-\alpha)} + (\widehat{\beta}_2(\omega) - \beta'_2)^2 \frac{[X^\top X]_{22}}{2\widehat{\sigma}^2(\omega) F^{-1}_{\mathcal{F}^2_{n-2}}(1-\alpha)}.\end{aligned}$$

Le résultat est que l'aire de l'ellipse de confiance est plus petite que celle du rectangle, autrement dit, que l'ellipse est une région de confiance plus précise que le rectangle. Cela était attendu puisque l'ellipse est de niveau exactement $1-\alpha$ tandis que le rectangle est de niveau au moins $1-\alpha$ (par application de la borne de l'union).