UNS
Master MathMods

Année 2014-2015
Yannick Baraud

Test of Mathematical Statistics

1 hour

Name (prénom): **Surname (nom):**

1. Give the assumptions for Hoeffding's inequality and recall the inequality.
Donner les hypothèses de l'inégalité de Hoeffding et rappeler l'inégalité.

Let $X_1, X_2, \dots X_n$ be independent random variables
and $P_r(X_i \in [a_i, b_i]) = 1 \quad 1 \leq i \leq n$.

then

$$P_r\left(|\bar{X}_n - E[\bar{X}_n]| \geq t\right) \leq 2\exp \frac{-2t^2n^2}{\sum_{i=1}^n (b_i - a_i)^2}$$

for $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$

or

$$P_r\left(|S_n - E[S_n]| \geq t\right) \leq 2\exp \frac{-2t^2}{\sum_{i=1}^n (b_i - a_i)^2}$$

for $S_n = \sum_{i=1}^n X_i$

Let X be distributed uniformly over $[0, \theta]$ for some $\theta > 0$, that is, X admits a density $x \mapsto f_\theta(x) = \theta^{-1} 1_{[0,\theta]}(x)$. In the sequel, we consider a n-sample $X_1, \ldots, X_n$ with the same distribution as X.
Soit X uniformément distribué sur $[0, \theta]$ pour un certain $\theta \in [0, 1]$, c'est à dire que X admet une densité $x \mapsto f_\theta(x) = \theta^{-1} 1_{[0,\theta]}(x)$. Dans la suite, on considère des variables aléatoires $X_1, \ldots, X_n$ de même loi que X.

2. Compute $\mathbb{E}_\theta\left[X^k\right]$ for all $k \geq 1$.
Calculer $\mathbb{E}_\theta\left[X^k\right]$ pour tout $k \geq 1$.

$$E_\theta[X^k] = \int_{-\infty}^{\infty} x^k f(x)\,dx = \frac{1}{\theta}\int_0^\theta x^k\,dx = \frac{1}{\theta}\left.\frac{x^{k+1}}{k+1}\right|_0^\theta$$

$$= \frac{\theta^k}{k+1}$$

3. On the basis of $\mathbb{E}_\theta[X]$, compute the estimator $\widehat{\theta}_{n,1}$ of θ using the method of moments.
En utilisant $\mathbb{E}_\theta[X]$, calculer l'estimateur $\widehat{\theta}_{n,1}$ de θ basé sur la méthode des moments.

$$E_\theta[X] = \frac{\theta}{2}, \qquad \frac{\hat\theta_{n,1}}{2} = \frac{1}{n}\sum_{i=1}^n X_i$$

$$\hat\theta_{n,1} = \frac{2}{n}\sum_{i=1}^n X_i$$

Method of Moments

$$\hat\mu = \frac{1}{n}\sum_1^n X_i$$

$$\hat\mu^2 + \hat\sigma^2 = \frac{1}{n}\sum_1^n X_i^2, \qquad \hat\sigma^2 = \frac{1}{n}\sum_1^n X_i^2 - \left(\frac{1}{n}\sum_1^n X_i\right)^2$$

$$\mathrm{var}(X) = \frac{\theta^2}{12}$$

4. Compute the quadratic risk of $\widehat{\theta}_{n,1}$.
   Calculer le risque quadratique de $\widehat{\theta}_{n,1}$

$$R(\hat{\theta},\theta) = \mathrm{var}(\hat{\theta}) + \mathrm{Bias}^2(\hat{\theta})$$

Bias$(\hat{\theta})$

$$E[\hat{\theta}] - \theta$$
$$E[2\bar{X}_n] - \theta$$
$$\frac{2}{n} E[\sum X_i] - \theta$$
$$\frac{2}{n} n E[X_i] - \theta$$
$$2\frac{\theta}{2} - \theta = 0$$

unbiased

Var$(\hat{\theta})$

$$\mathrm{Var}\left(\frac{2}{n}\sum_1^n X_i\right)$$
$$= \frac{4}{n^2}\mathrm{var}\left(\sum_1^n X_i\right) \overset{\text{iid}}{=} \frac{4}{n^2} n \,\mathrm{Var}(X)$$
$$= \frac{4}{n}\left[E[X^2] - (E[X])^2\right]$$
$$= \frac{4}{n}\left[\frac{\theta^2}{3} - \left(\frac{\theta}{2}\right)^2\right] = \frac{4}{n}\frac{\theta^2}{12} = \frac{\theta^2}{3n}$$

$$R(\hat{\theta},\theta) = \frac{\theta^2}{3n}$$

5. Is $\widehat{\theta}_{n,1}$ asymptotically normal? Compute the limit of
   L'estimateur $\widehat{\theta}_{n,1}$ *est-il asymptotiquement normal? Calculer la limite de*

$$\sqrt{n}\left(\widehat{\theta}_{n,1} - \theta\right).$$

X_i iid

By CLT

$$\sigma = \mathrm{var}(\hat{\theta}_{n,2}) = \frac{\theta^2}{3n}.$$

$$\frac{\sqrt{n}}{\sigma}\left(\frac{1}{n}\sum^n X_i - E[X]\right) \xrightarrow{\mathcal{L}} N(0,1)$$

$$\sqrt{n}\left(\frac{1}{n}\sum^n X_i - E[X]\right) \xrightarrow{\mathcal{L}} N(0, \mathrm{var}(X))$$

$$\sqrt{n}\left(\frac{2}{n}\sum^n X_i - 2E[X]\right) \xrightarrow{\mathcal{L}} 2N(0, \mathrm{var}(X))$$

$$\sqrt{n}\left(\hat{\theta}_{n,1} - \cancel{2}\frac{\theta}{\cancel{2}}\right) \xrightarrow{\mathcal{L}} N\left(0, 4\left(\frac{\theta^2}{12}\right)\right)$$

$$\sqrt{n}\left(\hat{\theta}_{n,1} - \theta\right) \xrightarrow{\mathcal{L}} N\left(0, \frac{\theta^2}{3}\right)$$

6. On the basis of $\mathbb{E}_\theta[X^2]$, compute the estimator $\widehat{\theta}_{n,2}$ of θ using the method of moments.
   En utilisant $\mathbb{E}_\theta[X^2]$, calculer l'estimateur $\widehat{\theta}_{n,2}$ de θ basé sur la méthode des moments.

$E[X^2] = \frac{\theta^2}{3}$

$\hat{\mu}^2 + \hat{\sigma}^2 = \frac{1}{n}\sum_i^n X_i^2 = \frac{\hat{\theta}_n^2}{3}$

$\hat{\theta}_{n,2} = \sqrt{\frac{3}{n}\sum_i^n X_i^2}$

$Var(X) = E[X^2] - (E[X])^2$

$= \frac{\theta^2}{3} - \left(\frac{\theta}{2}\right)^2 = \frac{\theta^2}{12}$

$Var(\hat{\theta}_{n,1}) = \frac{\theta^2}{3}$

7. Compute the limit of
   Calculer la limite de

$$\sqrt{n}\left(\widehat{\theta}_{n,2} - \theta\right).$$

By CLT. X_i^2 iid

$\sqrt{n}\left(\frac{1}{n}\sum_i^n X_i^2 - E[X_i^2]\right) \xrightarrow{\mathcal{L}} N(0, var(x^2))$

Delta Method: $\sqrt{n}(g(X_n) - g(\theta)) \xrightarrow{\mathcal{L}} (g'(\theta))\, N(0, var(x))$

$\sqrt{n}\left(\frac{1}{n}\sum^n X_i^2 - \frac{\theta^2}{3}\right) \xrightarrow{\mathcal{L}} N\left(0, \frac{4}{45}\theta^4\right)$

$g(x) = \sqrt{3x}$ $\quad g'(x) = \frac{1}{2}\frac{3}{\sqrt{3x}}$

$\sqrt{n}\left(\sqrt{\frac{3}{n}\sum^n X_i^2} - \theta\right) \xrightarrow{\mathcal{L}} g'\left(\frac{\theta^2}{3}\right) N\left(0, \frac{4}{45}\theta^4\right)$

$\xrightarrow{\mathcal{L}} \frac{\sqrt{3}}{2\sqrt{\frac{\theta^2}{3}}} N\left(0, \frac{4}{45}\theta^4\right)$

$\xrightarrow{\mathcal{L}} \frac{3}{2\theta} N\left(0, \frac{4}{45}\theta^4\right)$

$\sqrt{n}\left(\hat{\theta}_{n,2} - \theta\right) \xrightarrow{\mathcal{L}} N\left(0, \frac{\theta^2}{5}\right)$

$Var(\hat{\theta}_{n,2}) = var\left(\frac{3}{n}\sum_i^n X_i^2\right)$

$= \frac{9}{n^2} var\left(\sum_i^n X_i^2\right)$

$= \frac{9}{n^2}\left[E[X^4] - (E[X^2])^2\right]$

$= \frac{9}{n^2}\left[\frac{1}{\theta}\int_0^\theta x^4\,dx - \left(\frac{\theta^2}{3}\right)^2\right]$

$= \frac{9}{n^2}\left[\frac{\theta^4}{5} - \frac{\theta^4}{9}\right] =$

$= \frac{9}{n^2}\frac{4\theta^4}{45} = \frac{4}{5n^2}\theta^4$

$Var(X_i^2) = \frac{n^2}{9} var(\hat{\theta}_{n,2})$

$= \frac{4}{45}\theta^4$

8. Using the previous limit, build an asymptotic confidence interval $\widehat{I_n}$ for θ with level $1-\alpha$.
En utilisant la limite précédente, construire un intervalle de confiance $\widehat{I_n}$ pour θ de niveau $1-\alpha$.

$$\sqrt{n}\left(\hat{\theta}_{n,2}-\theta\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\theta^2}{5}\right)$$

$$\sqrt{n}\left(\hat{\theta}_{n,2}-\theta\right) \xrightarrow{\mathcal{L}} \frac{\theta}{\sqrt{5}} \mathcal{N}(0,1)$$

$$\frac{\sqrt{n}\sqrt{5}}{\theta}\left(\hat{\theta}_{n,2}-\theta\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

$$\mathbb{P}\left(\left|\frac{\sqrt{5n}}{\theta}\left(\hat{\theta}_{n,2}-\theta\right)\right| \leq t_\alpha\right) \geq 1-\alpha$$

$$-t_\alpha \leq \sqrt{5n}\left(\frac{\hat{\theta}_{n,2}}{\theta}-1\right) \leq t_\alpha$$

$$1 - \frac{t_\alpha}{\sqrt{5n}} \leq \frac{\hat{\theta}_{n,2}}{\theta} \leq \frac{t_\alpha}{\sqrt{5n}} + 1$$

$$\frac{\hat{\theta}_{n,2}}{1+\frac{t_\alpha}{\sqrt{5n}}} \leq \theta \leq \frac{\hat{\theta}_{n,2}}{1-\frac{t_\alpha}{\sqrt{5n}}}$$

9. Compute the cumulative distribution function F of X and the median $m(\theta)$ of X.
Calculer la fonction de répartition F de X et la médiane $m(\theta)$ de X.

$$F(x) = \int_{-\infty}^{x} f(x)\,dx = \int_{-\infty}^{x} \frac{1}{\theta} \mathbb{1}_{0,\theta}\,dx = \int_{0}^{x} \frac{1}{\theta}\,dx = \frac{1}{\theta}x \qquad x \le \theta$$

$$F(x) = \frac{1}{2} \qquad x = F^{-1}\left(\frac{1}{2}\right) \qquad y = \frac{1}{\theta}x, \quad x = \theta y$$

$$x = \theta \frac{1}{2} = m(\theta)$$

$$m(\theta) = \frac{\theta}{2} = q\left(\frac{1}{2}\right)$$

x point s.t 50% data are below

10. Give an estimator $\widehat{m}_n$ of $m(\theta)$ based on the empirical quantiles and compute the limit
Donner un estimateur $\widehat{m}_n$ de $m(\theta)$ basé sur les quantiles empiriques et calculer la limite

$$\sqrt{n}\,(\widehat{m}_n - m(\theta)).$$

$F'(x) = f(x)$

Thm: $$\sqrt{n}\left(\hat{q}_n(u) - q(u)\right) \xrightarrow{L} N\left(0, \frac{u(1-u)}{(F'(q(u)))^2}\right)$$

$$\sqrt{n}\left(\hat{q}_n\left(\tfrac{1}{2}\right) - q\left(\tfrac{1}{2}\right)\right) \xrightarrow{L} N\left(0, \frac{\frac{1}{2}\left(1-\frac{1}{2}\right)}{\left(f\left(q\left(\frac{1}{2}\right)\right)\right)^2}\right)$$

$$\sqrt{n}\left(\hat{q}_n\left(\tfrac{1}{2}\right) - q\left(\tfrac{1}{2}\right)\right) \xrightarrow{L} N\left(0, \frac{1/4}{\left(f\left(\frac{\theta}{2}\right)\right)^2}\right)$$

$$f\left(\frac{\theta}{2}\right) = \frac{1}{\theta}$$

$$\xrightarrow{L} N\left(0, \frac{\theta^2}{4}\right)$$

11. Using the limit above, build from this result an asymptotic confidence interval $\widetilde{I}_n$ for θ with level $1-\alpha$.
En utilisant la limite précédente, construire un intervalle de confiance $\widetilde{I}_n$ pour θ de niveau $1-\alpha$.

as $m = q(\frac{1}{2}) = \frac{\theta}{2}$. $\quad \sqrt{n}(\hat{m}_n - \hat{m}) \xrightarrow{\mathcal{L}} N(0, \frac{\theta^2}{4})$

$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} 2N(0, \theta^2/4) \Rightarrow \frac{\sqrt{n}}{\theta}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} N(0,1)$

$-t_\alpha \leq \frac{\sqrt{n}}{\theta}(\hat{\theta}_n - \theta) \leq t_\alpha \Rightarrow 1 - \frac{t_\alpha}{\sqrt{n}} \leq \frac{\hat{\theta}_n}{\theta} \leq \frac{t_\alpha}{\sqrt{n}} + 1$

$\frac{\hat{\theta}_n}{1 + t_\alpha/\sqrt{n}} \leq \theta \leq \frac{\hat{\theta}_n}{1 - t_\alpha/\sqrt{n}}$

12. Between $\widehat{I}_n$ and $\widetilde{I}_n$, what confidence interval would you use? explain why.
Entre les intervalles $\widehat{I}_n$ et $\widetilde{I}_n$, quel intervalle utiliseriez-vous? Expliquer pourquoi.

UNS
Master MathMods/MPA

Année 2014-2015
Yannick Baraud

Mathematical Statistics
3 hours

1. Give the Central Limit Theorem (assumptions and result).
   Donner le théorème de la limite centrale (hypothèses et conclusions)

2. Give the law of large numbers (assumptions and result)
   Donner la loi des grands nombres (hypothèses et conclusion).

3. Give Cochran's Theorem.
   Donner le théorème de Cochran

Exercise 1 Let $X_1, \dots, X_n$ be i.i.d. random variables with uniform distribution on $[0, \theta]$ with $\theta > 0$.
Soit $X_1, \dots, X_n$ i.i.d. de loi uniforme sur $[0, \theta]$ avec $\theta > 0$.

4. What is the maximum likelihood estimator $\widehat{\theta}_n$ of θ?
   Quel est l'estimateur du maximum de vraisemblance $\widehat{\theta}_n$ de θ?

5. Compute the cumulative distribution function of $\widehat{\theta}_n$.
   Calculer la fonction de répartition de $\widehat{\theta}_n$

6. Compute the density of $\widehat{\theta}_n$.
   Calculer la densité de $\widehat{\theta}_n$

7. For which value of $a \in \mathbb{R}$, $a\widehat{\theta}_n$ is unbiased.
   Pour quelle valeur de $a \in \mathbb{R}$, $a\widehat{\theta}_n$ est-il sans biais?

8. For which value of $b \in \mathbb{R}$, the quadratic risk of $b\widehat{\theta}_n$ is minimum?
   Pour quelle valeur de $b \in \mathbb{R}$, le risque quadratique de $b\widehat{\theta}_n$ est-il minimum?

Exercise 2 Let $X_1, \dots, X_n$ be i.i.d. Gaussian random variables with mean $\theta \in \mathbb{R} = \Theta$ and variance 1.
soit $X_1, \dots, X_n$ i.i.d. gaussiènne de moyenne $\theta \in \mathbb{R} = \Theta$ et de variance 1.

9. Give an estimator $\widehat{\theta}_n$ of θ based on the moment method.
   Donner un estimator $\widehat{\theta}_n$ de θ basé sur la méthode des moments.

10. Give the estimator $\widetilde{\theta}_n$ of θ based on the median.
    Donner un estimator $\widetilde{\theta}_n$ de θ basé sur la médiane.

11. Show that $\sqrt{n}\left(\widetilde{\theta}_n - \theta\right)$ converges to some limit distribution and specify the limit.
Montrer que $\sqrt{n}\left(\widetilde{\theta}_n - \theta\right)$ *converge en loi et préciser la limite.*

We now use a Bayesian approach to estimate θ and use the prior ν_τ on Θ given by the Gaussian distribution with mean 0 and variance $\tau^2 > 0$.
On utilise à présent une approche bayésienne pour estimer θ *basée sur la loi a priori* ν_τ *qui est gaussiènne de moyenne 0 et de variance* $\tau^2 > 0$

12. Show that the posterior distribution is Gaussian with mean $\frac{n\tau^2}{n\tau^2+1}\overline{X}_n$ and variance $\frac{\tau^2}{n\tau^2+1}$
Montrer que la loi a posteriori est gaussiènne de moyenne $\frac{n\tau^2}{n\tau^2+1}\overline{X}_n$ *et de variance* $\frac{\tau^2}{n\tau^2+1}$

13. Compute the Bayes estimator $\widehat{\theta}_n^\tau$ associated to the prior ν_τ.
Calculer l'estimateur bayésien $\widehat{\theta}_n^\tau$ *associé à la loi a priori* ν_τ

14. Compute the quadratic risk $R(\theta, \overline{X}_n)$ of $\overline{X}_n$ at θ.
Calculer le risque quadratique $R(\theta, \overline{X}_n)$ *de* $\overline{X}_n$ *en* θ

15. Compute the quadratic risk $R(\theta, \widehat{\theta}_n^\tau)$ of $\widehat{\theta}_n^\tau$ at θ.
Calculer le risque quadratique $R(\theta, \widehat{\theta}_n^\tau)$ *de* $\widehat{\theta}_n^\tau$ *en* θ

Given an estimator $\overline{\theta}_n$ of θ, we denote by $R^\tau(\overline{\theta}_n)$, the Bayes risk of $\overline{\theta}_n$ with respect to the prior ν_τ.
Etant donné un estimateur $\overline{\theta}_n$ *de* θ, *on note* $R^\tau(\overline{\theta}_n)$, *the risk de Bayes de* $\overline{\theta}_n$ *associé à la loi a priori* ν_τ

16. Compute $R_n = R^\tau(\overline{X}_n)$ and show that this quantity is independent of τ.
Calculer $R_n = R^\tau(\overline{X}_n)$ *et montrer que cette quantité est indépendante de* τ.

17. Compute $R^\tau(\widehat{\theta}_n^\tau)$
Calculer $R^\tau(\widehat{\theta}_n^\tau)$

18. Show that
Montrer que
$$\lim_{\tau \to +\infty} R^\tau(\widehat{\theta}_n^\tau) = \sup_{\theta \in \Theta} R(\theta, \overline{X}_n)$$

19. Show that $\overline{X}_n$ is minimax.
Montrer que $\overline{X}_n$ *est minimax.*

Exercise 3 Let $X_1, \ldots, X_n$ be i.i.d. random variables with density
Soit $X_1, \ldots, X_n$ i.i.d. de densité

$$f_\theta(x) = \frac{\theta}{x^{\theta+1}} \mathbb{1}_{x \geq 1} \quad \text{with} \quad \theta \in \Theta = (0, +\infty).$$

20. Compute the cumulative distribution function of X_1.
Calculer la fonction de répartition de X_1

21. Show that the statistical model is an exponential family.
Montrer que le modèle statistique est une famille exponentielle

22. What is the maximum likelihood $\widehat{\theta}_n$ estimator of θ?
Quel est l'estimateur du maximum de vraisemblance?

23. What is the limit in distribution of $\sqrt{n}\left(\widehat{\theta}_n - \theta\right)$?
Quelle est la limite en loi de $\sqrt{n}\left(\widehat{\theta}_n - \theta\right)$?

24. Build an asymptotic two-sided confidence interval for θ.
Construire un intervalle de confiance asymptotique bilatère pour θ