MAP 311: Aléatoire PC 8

Marc Lelarge

20 juin 2016

Exercice 1

Soit $(X_n)_{n\geq 1}$ une suite de v.a. indépendantes et de même loi. On suppose que X_1 est de carré intégrable, de moyenne m et de variance $\sigma^2 > 0$. On définit:

$$\hat{m}_n = \frac{X_1 + \cdots + X_n}{n}, \quad \hat{\sigma}_n^2 = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \hat{m}_n)^2.$$

1. Jusitifier que $\sqrt{n}\frac{\hat{m}_n - m}{\sigma}$ converge en loi vers $\mathcal{N}(0,1)$.
2. En déduire un intervalle de confiance asymptotique pour m au niveau 95% (en supposant σ connu).
3. Montrer que $\hat{\sigma}_n$ converge presque sûrement vers σ. L'estimateur $\hat{\sigma}_n^2$ est-il sans biais? On pourra montrer que $\frac{n-1}{n}\hat{\sigma}_n^2 = \frac{1}{n}\sum_{k=1}^n X_k^2 - \hat{m}_n^2$.
4. Montrer que
$$\sqrt{n}\frac{\hat{m}_n - m}{\hat{\sigma}_n} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1).$$
5. En déduire un intervalle de confiance asymptotique pour m au niveau 95%.

Solution.

1. TCL.
2. Soit $z_{\alpha/2}$ le quantile d'ordre $1-\alpha/2$ de la loi normale $\mathbb{P}(Z \geq z_{\alpha/2}) = \alpha/2$. On a alors
$$\mathbb{P}\left(-z_{\alpha/2} \leq \sqrt{n}\frac{\hat{m}_n - m}{\sigma} \leq z_{\alpha/2}\right) \to \mathbb{P}(|Z| \leq z_{\alpha/2}) = 1 - \alpha.$$
L'intervalle $\left[\hat{m}_n - \frac{\sigma z_{\alpha/2}}{\sqrt{n}}, \hat{m}_n + \frac{\sigma z_{\alpha/2}}{\sqrt{n}}\right]$ est un intervalle de confiance asymptotique au niveau $1-\alpha$. Pour $\alpha = 0.005$, on prend $z_{\alpha/2} \approx 1.96$.
3. On a
$$\hat{\sigma}_n^2 = \frac{1}{n}\sum_{k=1}^n X_k^2 - 2\hat{m}_n^2 + \hat{m}_n^2 = \frac{1}{n}\sum_{k=1}^n X_k^2 - \hat{m}_n^2.$$
Par la loi forte des grands nombres, $\frac{1}{n}\sum_{k=1}^n X_k^2$ c.v. p.s. vers $\mathbb{E}[X_1^2] = \sigma^2 + m^2$ et $\hat{m}_n$ c.v. p.s. vers m donc $\hat{\sigma_n}$ c.v. p.s. vers σ.

   On a
$$\mathbb{E}[\hat{m}_n^2] = \frac{\sigma^2}{n} + m^2,$$

donc

$$\mathbb{E}\left[\frac{n-1}{n}\hat{\sigma}_n^2\right] = \sigma^2 + m^2 - \left(\sigma^2/n + m^2\right) = \frac{n-1}{n}\sigma^2,$$

donc $\hat{\sigma}_n^2$ est un estimateur sans biais de la variance.

4. Comme $\hat{\sigma}_n$ converge en probabilité vers σ, on obtient le résultat en appliquant le théorème de Slutsky.

5. Comme dans la question 2, l'intervalle $\left[\hat{m}_n - \frac{\hat{\sigma}_n z_{\alpha/2}}{\sqrt{n}}, \hat{m}_n + \frac{\hat{\sigma}_n z_{\alpha/2}}{\sqrt{n}}\right]$ est un intervalle de confiance asymptotique au niveau $1-\alpha$.

Exercice 2:
Le but de cet exercice est d'estimer un temps d'attente qui suit une loi exponentielle de paramètre $\lambda > 0$ inconnu. Soit $(E_i)_{i\geq 1}$ une suite de v.a.r. i.i.d. de loi exponentielle de paramètre $\lambda > 0$ inconnu. On pose $\hat{m}_n = \frac{E_1+\cdots+E_n}{n}$ et

$$\hat{\sigma}_n^2 = \frac{1}{n-1}\sum_{k=1}^{n}(E_k - \hat{m}_n)^2.$$

1. Donner un intervalle de confiance asymptotique pour $\frac{1}{\lambda}$ au niveau 95%. En déduire un intervalle de confiance asymptotique pour λ au niveau 95%

2. Appliquer la méthode delta avec la fonction $f(x) = 1/x$.

Solution.

1. Comme $\mathbb{E}[E_1] = \frac{1}{\lambda}$, par l'exercice précédent $\left[\hat{m}_n - \frac{\hat{\sigma}_n z_{\alpha/2}}{\sqrt{n}}, \hat{m}_n + \frac{\hat{\sigma}_n z_{\alpha/2}}{\sqrt{n}}\right]$ est un intervalle de confiance asymptotique pour $\frac{1}{\lambda}$ au niveau $1-\alpha$. On en déduit que

$$\left[\frac{1}{\hat{m}_n + \frac{\hat{\sigma}_n z_{\alpha/2}}{\sqrt{n}}}, \frac{1}{\hat{m}_n - \frac{\hat{\sigma}_n z_{\alpha/2}}{\sqrt{n}}}\right]$$

est un intervalle de confiance pour λ.

2. en appliquant la méthode delta avec la fonction $f(x) = 1/x$, on a $\sqrt{n}\,(f(\hat{m}_n) - f(1/\lambda))$ qui converge en loi vers $\mathcal{N}(0, (f'(1/\lambda))^2\sigma^2)$. Comme la variance de E_1 $\sigma^2 = \frac{1}{\lambda^2}$ et $f'(1/\lambda) = -\lambda^2$ pour $\lambda > 0$, on obtient:

$$\sigma\sqrt{n}\left(\frac{1}{\hat{m}_n} - \lambda\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1).$$

En appliquant le lemme de Slutsky, on en déduit que

$$\left[\frac{1}{\hat{m}_n} - \frac{z_{\alpha/2}}{\hat{\sigma}_n\sqrt{n}}, \frac{1}{\hat{m}_n} + \frac{z_{\alpha/2}}{\hat{\sigma}_n\sqrt{n}}\right]$$

est un intervalle de confiance asymptotique pour λ au niveau $1-\alpha$.

Exercice 3: Soit $(X_n)_{n\geq 1}$ des v.a.r. i.i.d. de loi uniforme sur $[0,\theta]$.

1. Étudier la convergence, le biais, la variance, l'écart quadratique moyen, et la fluctuation asymptotique de l'estimateur $m_n = 2(X_1 + \cdots + X_n)/n$ de θ. Construire un intervalle de confiance asymptotique;

2. Montrer que l'estimateur $M_n = \max(X_1, \ldots, X_n)$ de θ a pour densité $(n/\theta)(x/\theta)^{n-1}\mathbf{1}_{[0,\theta]}(x)$ puis calculer $\mathbb{E}(M_k)$ et $Var(M_k)$. Comparer avec l'estimateur moyenne empirique m_n;

3. Montrer que $M_n \xrightarrow{\mathbb{P}} \theta$. Montrer de deux manières différentes que la convergence est p.s.;

4. Montrer que $W_n = n(M_n/\theta - 1)$ converge en loi quand $n \to \infty$, et déterminer la limite. Construire un intervalle de confiance asymptotique.

5. Montrer que M_n est l'estimateur de maximum de vraisemblance de θ.

Solution. Pour un estimateur $\hat{\theta}_n$ de θ, l'écart quadratique moyen possède toujours une décomposition (carré-du-)biais–variance : $\mathbb{E}((\hat{\theta}_n - \theta)^2) = (\mathbb{E}(\hat{\theta}_n) - \theta)^2 + \mathrm{Var}(\hat{\theta}_n)$. Cela découle du théorème de Pythagore dans L^2: $\mathbb{E}((Z-\theta)^2) = \mathrm{Var}(Z) + (\mathbb{E}(Z) - \theta)^2$. Note: $\mathrm{Var}(Z) = \arg\min_{c\in\mathbb{R}} \mathbb{E}((Z-c)^2)$.

1. La LGN entraîne que $m_n \to \theta$ p.s. L'estimateur n'est pas biaisé : $\mathbb{E}(m_n) = \theta$. La variance est égale à l'écart quadratique moyen, et vaut $\frac{4}{n}\mathrm{Var}(X_1) = \frac{\theta^2}{3n} = \mathcal{O}(\frac{1}{n})$. La fluctuation asymptotique est gaussienne, de vitesse $\sqrt{n}$, car par le TCL, $\sqrt{n}(m_n - \theta) \xrightarrow{\text{loi}} \mathcal{N}(0, \frac{\theta^2}{3})$. L'intervalle de confiance asymptotique s'obtient avec le résultat de fluctuation. Nul besoin du lemme de Slutsky car l'inversion en θ est facile ici: si $Z \sim \mathcal{N}(0,1)$ et $J \subset \mathbb{R}$,
$$\mathbb{P}(Z \in J) \approx \mathbb{P}\left(\frac{\sqrt{3n}}{\theta}(m_n - \theta) \in J\right) = \mathbb{P}(\theta \in m_n/(1 + (3n)^{-1/2}J))$$
Tout intervalle J tel que $\mathbb{P}(Z \in J) = 1 - \alpha$ fournit l'intervalle de confiance $I = m_n/(1 + (3n)^{-1/2}J)$ de niveau $1-\alpha$ pour θ. On peut chercher à choisir J de sorte que I soit petit;

2. $F_{M_n}(x) = (x/\theta)^n\mathbf{1}_{[0,\theta]} + \mathbf{1}_{]\theta,\infty[}$ et $f_{M_n}(x) = F'_{M_n}(x) = (n/\theta)(x/\theta)^{n-1}\mathbf{1}_{[0,\theta]}(x)$. Par conséquent, $\mathbb{E}(M_n) = \frac{n}{n+1}\theta$ et $\mathrm{Var}(M_n) = \frac{n}{(n+1)^2(n+2)}\theta^2 = \mathcal{O}(\frac{1}{n^2})$. L'estimateur M_n est baisé (et asymptotiquement sans biais), mais il est plus rapide que m_n. L'écart quadratique moyen de M_n vaut $(\mathbb{E}(M_n) - \theta)^2 + \mathrm{Var}(M_n) = \frac{\theta^2}{(n+1)^2} + \frac{n\theta^2}{(n+1)^2(n+2)} = \frac{2}{(n+1)(n+2)}\theta^2 = \mathcal{O}(\frac{1}{n^2})$;

3. Pour tout $\varepsilon > 0$ assez petit, $\mathbb{P}(|M_n - \theta| > \varepsilon) = \mathbb{P}(M_n < \theta - \varepsilon) = ((\theta - \varepsilon)/\theta)^n$. Cela donne la convergence en probabilité de M_n vers θ, et même la convergence p.s. en utilisant le lemme de Borel-Cantelli pour ε fixé (il faut ensuite poser $\varepsilon = 1/k$ et faire $\cap_k$). Alternativement, comme p.s. $(M_n)_{n\geq 1}$ est positive croissante et majorée par θ, elle converge p.s. et dans L^1 vers une v.a.r. $\leq \theta$ de moyenne $\lim_{n\to\infty} \mathbb{E}(M_n) = \theta$, qui est forcément égale p.s. à θ;

4. $F_{W_n}(x) = \mathbb{P}(M_n \leq \theta(1 + x/n))$. Ainsi, $F_{W_n}(x) \to 1$ si $x > 0$, tandis que si $x \leq 0$ alors, pour $n \gg 1$, $F_{W_n}(x) = (1 + x/n)^n \to e^x$. Ainsi, $W_n \to -W$ en loi quand $n \to \infty$, o $W \sim \mathrm{Exp}(1)$. L'intervalle de confiance asymptotique s'obtient avec la fluctuation: Pour tout $J \subset \mathbb{R}$, on a
$$\mathbb{P}(-W \in J) \approx \mathbb{P}(W_n \in J) = \mathbb{P}(\theta \in M_n/(1 + J/n)).$$
N'importe quel intervalle J tel que $\mathbb{P}(-W \in J) = 1 - \alpha$ fournit l'intervalle de confiance asymptotique $I = M_n/(1 + J/n)$ pour θ. On peut choisir J tel que I soit petit.

5. Pour tout réel $\theta \geq 0$, on a $L(\theta; X_1, \ldots, X_n) = \theta^{-n} \prod_{k=1}^n \mathbf{1}_{[0,\theta]}(X_k)$, qui vaut 0 si $\theta < M_n$ et qui décroit comme θ^{-n} si $\theta \geq M_n$. Donc $M_n = \arg\max_{\theta \geq 0} L(\theta; X_1, \ldots, X_n)$.

Exercice 4: Théorème de Cochran

Soit X un vecteur colonne aléatoire de $\mathbb{R}^n$ de loi $\mathcal{N}(m, \sigma^2 I_n)$, et $\mathbb{R}^n = E_1 \oplus \cdots \oplus E_p$ une décomposition de $\mathbb{R}^n$ en somme directe de p sous-espaces vectoriels orthogonaux de dimensions $d_1, \ldots, d_p$ avec $d_1 + \cdots + d_p = n$. Soit $\mathbf{P}_k$ la matrice du projecteur orthogonal sur E_k et $Y_k := \mathbf{P}_k X$ la projection orthogonale de X sur E_k.

1. Montrer que les projections $(Y_1, \ldots, Y_p)$ sont indépendantes et $Y_k \sim \mathcal{N}(\mathbf{P}_k m, \sigma^2 \mathbf{P}_k)$;

2. Monter que $\|Y_1 - \mathbf{P}_1 m\|^2, \ldots, \|Y_p - \mathbf{P}_p m\|^2$ sont indépendantes et $\sigma^{-2}\|Y_k - \mathbf{P}_k m\|^2 \sim \chi^2(d_k)$.

Solution.

1. On se ramène tout d'abord au cas où $m = 0$ par translation. Le vecteur aléatoire $Y = (Y_1, \ldots, Y_p)^\top$ de $\mathbb{R}^{np}$ s'écrit $Y = AX$ où A est la matrice de dimension $np \times n$

$$A = \begin{pmatrix} \mathbf{P}_1 \\ \vdots \\ \mathbf{P}_p \end{pmatrix}.$$

Il en découle que Y suit la loi $\mathcal{N}(0, \sigma^2 AA^\top)$. Pour tout $1 \leq i \leq p$, on a $\mathbf{P}_i = \mathbf{P}_i^\top = \mathbf{P}_i^2$. De plus, $\mathbf{P}_i \mathbf{P}_j = 0$ si $1 \leq i \neq j \leq p$ car $E_i \perp E_j$. Par conséquent, $AA^\top = \mathrm{Diag}(\mathbf{P}_1, \ldots, \mathbf{P}_p)$ est diagonale par blocs. Il en découle que $Y_1, \ldots, Y_p$ sont des vecteurs gaussiens indépendants avec $Y_k \sim \mathcal{N}(0, \sigma^2 \mathbf{P}_k)$ pour tout $1 \leq k \leq p$.

2. Les variables aléatoires $\|Y_1\|^2, \ldots, \|Y_p\|^2$ sont indépendantes. Il reste à déterminer leur loi. Pour tout $1 \leq k \leq p$, soit $B_k = \{e_{k,1}, \ldots, e_{k,d_k}\}$ une base orthonormée de E_k. La réunion $B_1 \cup \ldots \cup B_p$ constitue une base orthonormée de $\mathbb{R}^n$. Le vecteur X s'écrit dans cette base $X = Y_1 + \cdots + Y_p$ avec $Y_k = a_{k,1} e_{k,1} + \cdots + a_{k,d_k} e_{k,d_k}$ où $a_{k,i} = \langle X, e_{k,i} \rangle$. L'invariance par transformation orthogonale de la loi $\mathcal{N}(0, \sigma^2 \mathrm{I}_n)$ implique que les variables aléatoires $a_{k,i}$ sont indépendantes et de même loi $\mathcal{N}(0, \sigma^2)$. Il en découle que $\sigma^{-2}\|Y_k\|^2 = \sigma^{-2}(a_{k,1}^2 + \cdots + a_{k,d_k}^2) \sim \chi^2(d_k)$ pour tout $1 \leq k \leq p$, ce qui achève la preuve du théorème de Cochran.

Exercice 5: Échantillons gaussiens

Soit $(X_1, \ldots, X_n)$ un échantillon de loi $\mathcal{N}(m, \sigma^2)$. On lui associe la *moyenne empirique* et la *variance empirique* définies respectivement par

$$\overline{X}_n = \frac{1}{n} \sum_{k=1}^n X_k \qquad \text{et} \qquad \sigma_n^2 = \frac{1}{n-1} \sum_{k=1}^n (X_k - \overline{X}_n)^2.$$

Monter que les variables aléatoires $\overline{X}_n$ et σ_n^2 sont indépendantes avec

$$\overline{X}_n \sim \mathcal{N}\left(m, \frac{\sigma^2}{n}\right) \qquad \text{et} \qquad \frac{(n-1)}{\sigma^2} \sigma_n^2 \sim \chi^2(n-1).$$

Solution. Soit 1_n le vecteur de $\mathbb{R}^n$ dont toutes les coordonnées sont égales à 1, et soit $E_1 = \{a1_n;\ a \in \mathbb{R}\}$ le sous-espace vectoriel de $\mathbb{R}^n$ engendré par 1_n. La matrice de la projection orthogonale sur E_1 est donnée par

$$\mathbf{P}_1 = \frac{1_n 1_n^\top}{\|1_n\|^2} = \frac{1}{n} 1_n 1_n^\top.$$

Le sous-espace $E_2 = E_1^\perp$ est de dimension $n-1$ et la matrice de la projection orthogonale sur E_2 est $\mathbf{P}_2 = \mathrm{I}_n - \mathbf{P}_1$. On a $Y_1 = \mathbf{P}_1 X = \overline{X}_n 1_n$ et $Y_2 = \mathbf{P}_2 X = (X_1 - \overline{X}_n, \ldots, X_n - \overline{X}_n)^\top$, ce qui entraîne $\|Y_2\|^2 = (n-1)\sigma_n^2$. Le théorème de Cochran permet de conclure.

Exercice 6: Test d'aquéquation du chi-deux

Il s'agit d'un test non-paramétrique classique.

1. Soient $p = (p_1, \ldots, p_k)$, $q = (q_1, \ldots, q_k)$ des lois discrètes avec $p_i > 0$ pour tout $1 \le i \le k$, et

$$D(p,q) := \sum_{i=1}^k \frac{(p_i - q_i)^2}{p_i}.$$

   Montrer que $D(p,q) \ge 0$ avec égalité ssi $p = q$. S'agit-il d'une distance?

2. Soit $X_1, \ldots, X_n$ un échantillon d'une loi discrète inconnue $q = (q_1, \ldots, q_k)$ sur $\{1, \ldots, k\}$. On souhaite savoir si cette loi q est égale à une loi discrète de référence $p = (p_1, \ldots, p_k)$ (qui est connue). Pour cela, on introduit les hypothèses statistiques antagonistes:

$$H_0 : p = q, \quad H_1 : p \neq q.$$

   On introduit la statistique de test $S_n = S(X_1, \ldots, X_n)$ définie par

$$S_n := nD(p, \widehat{q}) = n \sum_{i=1}^k \frac{(p_i - \widehat{q}_i)^2}{p_i} = n \sum_{i=1}^k \frac{(n_i - N_i)^2}{n_i}$$

   où

   - $\widehat{q} = (N_1/n, \ldots, N_k/n)$ est l'estimateur empirique de q;
   - $N_i := \mathrm{card}\{1 \le m \le n : X_m = i\} = \sum_{m=1}^n \mathbf{1}_{\{X_m = i\}}$ est l'effectif de i dans l'échantillon;
   - $n_i := np_i$ est l'effectif théorique sous H_0.

   En pratique, on calcule les n_i et les N_i (comptage) puis S_n. Montrer que

   - si H_0 est fausse alors $S_n \to +\infty$ presque sûrement quand $n \to \infty$;
   - si H_0 est vraie alors $S_n \to \chi^2(k-1)$ en loi quand $n \to \infty$.

   Indication: établir que

$$S_n = \left\| \frac{1}{\sqrt{n}} \sum_{m=1}^n V_m \right\|_2^2 \quad \text{où} \quad V_{m,i} := \frac{\mathbf{1}_{\{X_m = i\}} - p_i}{\sqrt{p_i}}.$$

3. Fixons un seuil de tolérance $0 < \alpha < 1$, typiquement $\alpha = 5/100$, appelé *niveau du test*, et considérons la région $\mathcal{R}_\alpha = [0, a]$ où a est le quantile $1-\alpha$ de la loi $\chi^2(k-1)$, appelée *région d'acceptation du test*. Au vu de $X_1, \ldots, X_n$, on décide comme suit:
   - Si $S_n \in \mathcal{R}_\alpha$, on accepte l'hypothèse H_0;
   - Si $S_n \notin \mathcal{R}_\alpha$, on rejette l'hypothèse H_0.

   De manière résumée la décision prise avec le test est $T_n := H_{\mathbf{1}_{S_n \notin \mathcal{R}_\alpha}}$. Montrer que
   - Si H_0 est vraie alors la probabilité de rejeter à tort H_0 tend vers α quand $n \to \infty$;
   - Si H_0 est fausse, alors la probabilité d'accepter à tort H_0 tend vers 0 quand $n \to \infty$.

   On parle d'*erreurs de première et de seconde espèce du test* ($\to$ niveau et puissance).

Solution.

1. Évident. Ce n'est pas une distance, car elle n'est pas symétrique. Cependant, $\sqrt{D(p,q)}$ est la distance euclidienne de p q pondérée par p;

2. D'après la loi forte des grands nombres,
$$\frac{S_n}{n} \xrightarrow[n\to\infty]{\text{p.s.}} \sum_{i=1}^{k} \frac{(p_i - q_i)^2}{p_i} = D(p,q).$$
   Si H_0 est fausse alors $D(p,q) > 0$ et donc $S_n = n(S_n/n) \sim_{n\to\infty} nD(p,q) \to +\infty$ presque sûrement quand $n \to \infty$. Si H_0 est vraie, alors $q = p$ et $D(p,q) = 0$, et comportement de S_n quand $n \to \infty$ peut-être décrit par le théorème central limite multivarié. On a
$$S_n = n \sum_{i=1}^{k} \frac{\left(p_i - \frac{1}{n}\sum_{m=1}^{n} \mathbf{1}_{\{X_m = i\}}\right)^2}{p_i} = \left\| \frac{V_1 + \cdots + V_n}{\sqrt{n}} \right\|_2^2$$
   où $V_1, \ldots, V_n$ sont les vecteurs aléatoires de $\mathbb{R}^k$ définis par
$$V_{m,i} := \frac{1}{\sqrt{p_i}}(\mathbf{1}_{\{X_m=i\}} - p_i)$$
   pour tout $1 \le m \le n$ et $1 \le i \le k$. Or les vecteurs aléatoires $V_1, \ldots, V_n$ sont i.i.d. de matrice de covariance $\Sigma = I_k - \sqrt{p}\sqrt{p}^\top$ où $\sqrt{p}^\top := (\sqrt{p_1}, \ldots, \sqrt{p_k})$. Par conséquent, le TCL donne
$$\frac{1}{\sqrt{n}} \sum_{m=1}^{n} V_m \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0, \Sigma).$$
   Comme la norme est continue, on obtient que
$$S_n = \left\| \frac{1}{\sqrt{n}} \sum_{m=1}^{n} V_m \right\|_2^2 \xrightarrow[n\to\infty]{\text{loi}} \|Z\|_2^2$$
   où $Z \sim \mathcal{N}(0,\Sigma)$. La matrice Σ est de rang $k-1$. La matrice $\sqrt{p}\sqrt{p}^\top$ est la matrice de projection orthogonale sur $\mathrm{Vect}(\sqrt{p})$, tandis que $\Sigma = I_k - \sqrt{p}\sqrt{p}^\top$ est la matrice de projection orthogonale sur $\mathrm{vect}(\sqrt{p})^\perp$ dans $\mathbb{R}^k$. Le théorème de Cochran donne $Z \sim \chi^2(k-1)$.

3. Si H_0 est vraie alors $\mathbb{P}(T_n = 1) = \mathbb{P}(S_n > a) = 1 - \mathbb{P}(S_n \le a) \to 1 - (1-\alpha) = \alpha$ car $S_n \to \chi^2(k-1)$. Si H_0 est fausse, alors $\mathbb{P}(T_n = 0) = \mathbb{P}(S_n \le a) \to 0$ car $S_n \to +\infty$ p.s.