1 Fondements probabilistes

Définition de la fonction $\Gamma(z)$. $\forall z > 0, \Gamma(z) = \int_0^\infty x^{z-1}e^x \, dx$.

Propriétés de la fonction $\Gamma(z)$.

- $\forall z > 0, \Gamma(z) = (z-1)\Gamma(z-1)$
- $\forall z \in \mathbb{N}^*, \Gamma(z) = (z-1)!$
- $\Gamma(1) = 1$
- $\Gamma(1/2) = \sqrt{\pi}$

Définition de la fonction bêta. $\forall x, y < 0, \beta(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$.

Fonction de répartition marginale. Soient $X_1, \ldots, X_k$, avec $j < k$. Alors, $F_{1\ldots,j}(x_1,\ldots,x_j) = F_{\mathbf{X}}(x_1,\ldots,x_j,\infty,\ldots,\infty)$. Remplacer F par f pour la fonction de densité marginale.

Espérance d'une fonction d'une v.a. $\mathbb{E}\left[g(X)\right] = \begin{cases} \sum_x g(x)p_X(x) \text{ si } X \text{ est discrète} \\ \int_{\mathbb{R}} g(x)f_X(x) \quad dx \text{ si } X \text{ est continue} \end{cases}$.

Inégalité de Jensen. Soit g une fonction convexe. Alors, $g(\mathbb{E}\left[X\right]) < \mathbb{E}\left[g(X)\right]$. Inverser le sens de l'inégalité pour les fonctions concaves. Il y a égalité si $g(u) = u$.

Fonction génératrice des moments (f.g.m.) $M_X(t) = \mathbb{E}\left[e^{tX}\right] \; \forall t \; \in \mathbb{R}$ t.q. l'espérance existe.

Propriété de la f.g.m. $\mathbb{E}\left[X^k\right] = \frac{d^k}{dt^k}M_X(t)\Big|_{t=0}$

f.g.m. des moments conjoints. $M(t_1,\ldots,t_k) = \mathbb{E}\left[\exp(t_1X_1 + \cdots + t_kX_k)\right]$.

Propriété de la f.g.m. des moments conjoints. $\mathbb{E}\left[X_1^{j_1},\ldots,X_k^{j_k}\right] = \frac{\partial^{j_1+\ldots j_k}}{\partial t_1^{j_1}\ldots\partial t_k^{j_k}}M(t_1,\ldots,t_k)\Big|_{t_1=0,\ldots,t_k=0}$

Indépendance de v.a. continues. $X_1,\ldots X_k$ sont des v.a. indépendantes si $f(x_1,\ldots,x_k) = f_1(x_1)\ldots f_k(x_k) \;\; \forall\, (x_1,\ldots,x_k) \in \mathbb{R}^k$.

Conséquences de l'indépendance. (Remarque : les réciproques ne sont pas vraies !)

- $\mathbb{E}\left[g_1(X_1)\ldots g_k(X_k)\right] = \mathbb{E}\left[g_1(X_1)\right]\ldots\mathbb{E}\left[g_k(X_k)\right]$
- $\mathbb{C}\text{ov}(X_i,X_j) = 0 \; \forall \; i \neq j$
- $M(t_1,\ldots,t_k) = M_{X_1}(t_1)\ldots M_{X_k}(t_k)$

Espérance conditionnelle/itérée. $\mathbb{E}\left[X\right] = \mathbb{E}_Y\left\{\mathbb{E}_{X|Y}[X|Y]\right\}$

Variance conditionnelle/itérée. $\mathbb{V}\left[X\right] = \mathbb{V}_Y\{\mathbb{E}_{X|Y}[X|Y]\} + \mathbb{E}_Y\left\{\mathbb{V}_{X|Y}(X|Y)\right\}$

Espérance de variables aléatoires multidimensionnelles. $\mathbb{E}\left[\mathbf{AX}+\mathbf{B}\right] = \mathbf{A}\mathbb{E}\left[\mathbf{X}\right] + \mathbf{B}$.

Variance de variables aléatoires multidimensionnelles. $\mathbb{V}\left[\mathbf{AX}+\mathbf{B}\right] = \mathbf{A}\mathbb{V}\left[\mathbf{X}\right]\mathbf{A}'$.

Transformation de v.a. : Fonction de répartition. Soit X une v.a.r. avec loi connue. Soit $Y = h(X)$. On cherche $f_Y(y)$. $\mathbb{P}\left[Y \leq y\right] = \mathbb{P}\left[h(X) \leq y\right] = \mathbb{P}\left[X \leq A_y\right]$, où $A_y = \{x : h(x) \leq y\}$. On calcule ensuite $f(y) = F'(y)$.

Note pour les transformations de variables aléatoires : vérifier le domaine des variables.

Transformation de v.a. : Jacobien. Soient $\mathbf{X} = (X_1, \ldots, X_k)'$ et $Y_i = h_i(\mathbf{X}), i \in \{1, \ldots, k\}$, avec h_i des fonctions **bijective**. Soient $X_i = g_i(\mathbf{Y})$ (en pratique, choisir intelligemment les fonctions g_i). Alors, $f_Y(y) = f_X[g(y)]\,|\det(J)|\mathbb{1}_{R_y}(y)$, où $\det(J)$ est le déterminant du Jacobien :

$$J = \begin{pmatrix} \frac{\partial}{\partial y_1} g_1(y) & \cdots & \frac{\partial}{\partial y_k} g_1(y) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial y_1} g_k(y) & \cdots & \frac{\partial}{\partial y_k} g_k(y) \end{pmatrix}. \tag{1}$$

Transformation de v.a. : Jacobien. (k=1). g est une fonction **bijective**... $Y = g(X)$. Alors, $f_Y(y) = f_X[g^{-1}(y)] \left| \frac{d}{dy} g^{-1}(y) \right| \mathbb{1}_{R_{y}(y)}$.

Inégalité de Chebychev. (Souvent utilisée pour montrer la convergence en probabilité.) Soit X une v.a. de moyenne μ et de variance $\sigma^2 < \infty$. Alors, $\mathbb{P}[|X - \mu| > \varepsilon] \leq \sigma^2/n \;\forall\, \varepsilon > 0$.

Convergence en probabilité. $X_n \xrightarrow{P} X \iff \text{plim}_{n\to\infty} X_n = X \iff \lim_{n\to\infty} \mathbb{P}[|X_n - X| > \varepsilon] = 0,\ \forall \varepsilon > 0$.

Loi faible des grands nombres. $X_1, \ldots, X_n$ i.i.d. de moyenne μ et de variance $\sigma^2 < \infty$. Soit $\bar{X}_n = (X_1 + \cdots + X_n)/n$. Alors, $\bar{X}_n \xrightarrow{P} \mu$.

Convergence en loi. $X_n \xrightarrow{L} X \iff \lim_{n\to\infty} \mathbb{P}[X_n \leq x] = F_X(x)$.

Théorème central limite. Soient $X_1, \ldots, X_n$ des v.a. iid de moyenne μ et de variance $\sigma^2 < \infty$. Alors, $\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{L} Z$, où $Z \sim \mathcal{N}(0, 1)$.

Théorème de la fonction continue. Soit g une fonction continue. $X_n \xrightarrow{P} \mathbf{X} \Rightarrow g(\mathbf{X}_n) \xrightarrow{P} g(\mathbf{X})$ et $g(\mathbf{X}_n) \xrightarrow{d} \mathbf{X} \Rightarrow g(\mathbf{X}_n) \xrightarrow{d} g(\mathbf{X})$.

Théorème de Slutsky. Si $X_n \xrightarrow{d} X$ et $Y_n \xrightarrow{P} c$, alors $X_n Y_n \xrightarrow{d} cX$.

2 Lois échantillonnales

Génération de nombres aléatoires. Soit $X \sim F_X(x)$ une v.a. continue, et $U \sim \mathcal{U}(0,1)$. On simule une v.a. de distribution $F(x)$ en utilisant tirant des observations u, puis $x = F_X^{-1}(u)$.

Théorème. Soient $X_1, \ldots, X_n$ des variables aléatoires indépendantes avec $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$. Alors, $Y = a_0 + \sum_i a_i X_i \sim \mathcal{N}(a_0 + \sum_i a_i \mu_i, \sum_i a_i^2 \sigma_i^2)$.

2.1 Relations entre les différentes lois statistiques

Proposition. Soient $Z \sim \mathcal{N}(0,1)$ et $X = Z^2$. Alors, $X \sim \chi_1^2$.

Théorème. Soient $U \sim \chi_u^2$ et $V \sim \chi_v^2$. Alors, $U + V \sim \chi_{u+v}^2$.

Exemple. Soient $X_1, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$ et $S_*^2 = \frac{1}{n}\sum_i (X_i - \mu)^2$. Alors, $nS_*^2/\sigma^2 \sim \chi_n^2$.

Proposition. Soient $W \perp V$ telles que $W \sim \mathcal{N}(0,1)$ et $V \sim \chi_r^2$. Alors, $T = W/\sqrt{V/r} \sim t_r$.

Théorème. Soient $X_1, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$. Alors,

1. $\bar{X}_n \sim \mathcal{N}(\mu, \sigma^2/n)$
2. $(n-1)S_n^2/\sigma^2 \sim \chi_{n-1}^2$
3. $\bar{X}_n \perp S_n^2$

Corollaire. Soient $X_1, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$. Alors, $T = \frac{\bar{X}-\mu}{\sqrt{S_n^2/n}} \sim t_{n-1}$.

Proposition. Soient $U \perp V$ telles que $U \sim \chi_n^2$ et $V \sim \chi_m^2$. Alors, $W = \frac{U/n}{V/m} \sim F_{n,m}$.

Théorème. Soient $X_1, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(\mu, \sigma_X^2)$ et $Y_1, \ldots, Y_n \overset{iid}{\sim} \mathcal{N}(\mu, \sigma_Y^2)$ deux échantillons indépendants. Soient S_X^2 et S_Y^2 leur variance échantillonnale respective. Alors, $F = \frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2} \sim F_{n-1,m-1}$.

2.2 Autres résultats.

Proposition. Soient $\boldsymbol{X} = (X_1, \ldots, X_n)'$ un vecteur de variables aléatoire (iid) issues d'une loi quelconque avec $\mathbb{E}[X_i] = \mu, \mathbb{V}[X_i] = \sigma^2, m_3^* = \mathbb{E}\left[(X_i - \mu)^3\right] < \infty, m_4^* = \mathbb{E}\left[(X_i - \mu)^4\right] < \infty$. Soient $V_n^2 = \frac{1}{n}\sum_i \left(X_i - \bar{X}_n\right)^2$ et $S_n^2 = \frac{1}{n-1}\sum_i \left(X_i - \bar{X}^2\right)^2$. Alors,

- $\mathbb{E}\left[\bar{X}_n\right] = \mu$ et $\mathbb{V}\left[\bar{X}_n = \sigma^2/n\right]$
- $\mathbb{E}\left[S_n^2\right] = \sigma^2$ et $\mathbb{V}\left[S_n^2\right] = m_4^*/n - \frac{n-3}{n(n-1)}\sigma^4$
- $\mathbb{E}\left[V_n^2\right] = \frac{n-1}{n}\sigma^2$ et $\mathbb{V}\left[V_n^2\right] = \frac{m_4^*(n-1)^2}{n^3} - \frac{\sigma^4(n-1)(n-3)}{n^2}$
- $\mathbb{C}\mathrm{ov}(\bar{X}_n, S_n^2) = m_3^*/n$
- $\mathbb{C}\mathrm{ov}(\bar{X}_n, V_n^2) = \frac{m_3^*(n-1)}{n^2}$

X	$f_X(x)$	$\mathbb{E}[X]$	$\mathbb{V}[X]$	$M_X(t)$
$\mathcal{N}(\mu,\sigma^2); \mu \in \mathbb{R}, \sigma^2 > 0$	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \mathbb{1}_{\{x\in\mathbb{R}\}}$	μ	σ^2	$\exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$
$\mathcal{N}_p(\mathbf{m},\sigma^2\mathbf{\Sigma}); \mathbf{m} \in \mathbb{R}^p, \sigma^2 > 0$	$\frac{1}{(2\pi\sigma^2)^{p/2}\lvert\mathbf{\Sigma}^{1/2}\rvert}\sigma^{-p} \exp\left(-\frac{(\boldsymbol{x}-\mathbf{m})'\mathbf{\Sigma}^{-1}(\boldsymbol{x}-\mathbf{m})}{2\sigma^2}\right) \mathbb{1}_{\{x\in\mathbb{R}^p\}}$	$\mathbf{m}$	$\sigma^2\mathbf{\Sigma}$	$\exp\left(\mathbf{m}'\boldsymbol{t} + \frac{1}{2}\sigma^2\boldsymbol{t}'\mathbf{\Sigma}\boldsymbol{t}\right)$
$t_r; r \in \mathbb{N}^*$	$\frac{\Gamma((r+1)/2)}{\Gamma(r/2)} \frac{1}{\sqrt{r\pi}} \frac{1}{(1+x^2/r)^{(r+1)/2}} \mathbb{1}_{\{x\in\mathbb{R}\}}$	0 si $r > 1$	$\frac{r}{r-2}$ si $r > 2$	–
Exponentielle$(\lambda); \lambda > 0$	$\lambda e^{-\lambda x}\mathbb{1}_{\{x>0\}}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda-t}; t < \lambda$
Gamma$(\alpha,\lambda); \alpha > 0, \lambda > 0$	$\frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}\mathbb{1}_{\{x>0\}}$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$	$\left(\frac{\lambda}{\lambda-t}\right)^\alpha; t < \lambda$
$\chi^2_k; k \in \mathbb{N}^*$	$\frac{1}{\Gamma(k/2)2^{k/2}} x^{(k/2)-1} e^{-x/2}\mathbb{1}_{\{x>0\}}$	k	$2k$	$(1-2t)^{-k/2}; t < 1/2$
$F_{n,m}; n \in \mathbb{N}^* \in m \in \mathbb{N}^*$	$\frac{\Gamma((n+m)/2)(n/m)^{n/2}}{\Gamma(n/2)\Gamma(m/2)} \frac{x^{n/2-1}}{(1+nx/m)^{(n+m)/2}} \mathbb{1}_{\{x>0\}}$	$\frac{m}{m-2}$ si $m > 2$	$\frac{2m^2(n+m-2)}{n(m-2)^2(m-4)}$ si $m > 4$	–
Beta$(\alpha,\beta); \alpha > 0, \beta > 0$	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}\mathbb{1}_{\{0<x<1\}}$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	–
$\mathcal{U}_{[a,b]}; (a,b) \in \mathbb{R}^2$	$\frac{1}{b-a}\mathbb{1}_{\{a<x<b\}}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{tb}-e^{ta}}{t(b-a)}$

Tableau 1 – Distribution de certaines lois continues

X	$p_X(x)$	$\mathbb{E}[X]$	$\mathbb{V}[X]$	$M_X(t)$
Bernoulli(p)	$p^x(1-p)^{1-x}\mathbb{1}_{\{x\in\{0,1\}\}}$	p	$p(1-p)$	pe^t+1-p
Binomiale(n,p)	$\binom{n}{x}(1-p)^{n-x}p^x\mathbb{1}_{\{x\in\mathbb{N}\}}$	np	$np(1-p)$	$(pe^t+1-p)^n$
Géométrique(p) (échec)	$p(1-p)^x\mathbb{1}_{\{x\in\mathbb{N}\}}$	$\frac{1}{p}-1$	$\frac{1-p}{p^2}$	$\frac{p}{1-(1-p)e^t}$
Géométrique(p) (succès)	$p(1-p)^{x-1}\mathbb{1}_{\{x\in\mathbb{N}^*\}}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$
Uniforme(m)	$\frac{1}{m}\mathbb{1}_{\{x\in\{1,\ldots,m\}\}}$	$\frac{m+1}{2}$	$\frac{m^2-1}{12}$	$\frac{e^t(e^{mt}-1)}{e^t-1}$
Poisson(λ)	$e^{-\lambda}\lambda^x/x!\mathbb{1}_{\{x\in\mathbb{N}\}}$	λ	λ	$\exp(\lambda(e^t-1))$

Tableau 2 – Distribution de certaines lois discrètes