EXAMEN STATISTIQUE - 1MF2E

Lundi 4 Avril 2022

Partiel sans document (Une feuille A4 recto-verso autorisée)

Exercice 1 : Estimation (11 points)

Un ensemble de n votants s'exprime pour 3 candidats dont deux d'entre eux ont la même probabilité θ d'être élus. La vraisemblance associée à ce modèle statistique est définie par une loi multinomiale d'expression $L(n_1, n_2, n_3; \theta) = P[N_1 = n_1, N_2 = n_2, N_3 = n_3] = \frac{n!}{n_1!n_2!n_3!} p_1^{n_1} p_2^{n_2} p_3^{n_3}$ où N_i est la variable aléatoire associée au nombre de voix obtenues pour le $i$ème candidat (avec $n_1 + n_2 + n_3 = n$) et $\boldsymbol{p} = (p_1, p_2, p_3) = (\theta, \theta, 1 - 2\theta)$ regroupe les probabilités que chaque candidat reçoive une voix, qui dépend d'un unique paramètre inconnu $\theta \in]0, 1/2[$. On admettra que pour un tel modèle statistique, la loi marginale de N_i est une loi binomiale $B(n, p_i)$ de moyenne $E[N_i] = np_i$ et de variance $\text{var}(N_i) = np_i(1 - p_i)$.

1. (2pts) Déterminer l'estimateur du maximum de vraisemblance du paramètre θ noté $\widehat{\theta}_{\text{MV}}$ en utilisant la vraisemblance $L(n_1, n_2, n_3; \theta)$ et montrer qu'il ne dépend que de n et de N_3 (on prendra soin de vérifier que cette vraisemblance admet un unique maximum global).
   En utilisant les valeurs de p_i données dans l'énoncé, on obtient la vraisemblance

$$L(n_1, n_2, n_3; \theta) = \theta^{n_1+n_2}(1 - 2\theta)^{n_3}.$$

   On sait qu'il est plus facile de travailler avec la log-vraisemblance définie par

$$\ln L(n_1, n_2, n_3; \theta) = (n_1 + n_2) \ln \theta + n_3 \ln(1 - 2\theta)$$

   qui admet pour dérivée

$$\frac{\partial \ln L(n_1, n_2, n_3; \theta)}{\partial \theta} = \frac{n_1 + n_2}{\theta} + \frac{-2n_3}{1 - 2\theta}.$$

   Les variations de cette dérivée sont définies par

$$\begin{aligned}
\frac{\partial \ln L(n_1, n_2, n_3; \theta)}{\partial \theta} \geq 0 \quad & \Leftrightarrow \quad \frac{n_1 + n_2}{\theta} + \frac{-2n_3}{1 - 2\theta} \geq 0 \\
& \Leftrightarrow \quad (n_1 + n_2)(1 - 2\theta) - 2n_3\theta \geq 0 \\
& \Leftrightarrow \quad n_1 + n_2 \geq 2n\theta \\
& \Leftrightarrow \quad \theta \leq \frac{n_1 + n_2}{2n},
\end{aligned}$$

   où on a utilisé le fait que $\theta < 1/2$ et $n = n_1 + n_2 + n_3$. Ceci permet de faire un tableau de variation qui indique que $\theta = \frac{n_1+n_2}{2n}$ est le maximum global unique de la vraisemblance, d'où

$$\widehat{\theta}_{\text{MV}} = \frac{N_1 + N_2}{2n} = \frac{n - N_3}{2n}.$$

2. (2pts) Déterminer la moyenne et la variance de l'estimateur $\widehat{\theta}_{\text{MV}}$. Est-il un estimateur sans biais et convergent du paramètre θ ?
   En utilisant le fait que N_1 et N_2 suivent des lois binomiales, on peut déterminer la moyenne de $\widehat{\theta}_{\text{MV}}$ comme suit
$$E\left[\widehat{\theta}_{\text{MV}}\right] = \frac{1}{2n}(E[N_1] + E[N_2]) = \frac{n\theta + n\theta}{2n} = \theta.$$
   $\widehat{\theta}_{\text{MV}}$ est donc un estimateur non biaisé de θ. De même, en utilisant $\text{var}\,[N_3] = 2n\theta(1-2\theta)$, la variance de $\widehat{\theta}_{\text{MV}}$ s'écrit
$$\text{var}\left[\widehat{\theta}_{\text{MV}}\right] = \frac{1}{4n^2}\text{var}N_3 = \frac{\theta(1-2\theta)}{2n}.$$
   Comme $E\left[\widehat{\theta}_{\text{MV}}\right] - \theta = 0$, on en déduit que $\widehat{\theta}_{\text{MV}}$ est un estimateur non biaisé du paramètre θ. Comme de plus la variance de $\widehat{\theta}_{\text{MV}}$ tend vers 0 lorsque $n \to \infty$, $\widehat{\theta}_{\text{MV}}$ est un estimateur convergent du paramètre θ.

3. (2pts) L'estimateur $\widehat{\theta}_{\text{MV}}$ est-il l'estimateur efficace du paramètre θ ?
   La dérivée seconde de la vraisemblance est
$$\frac{\partial^2 \ln L(n_1, n_2, n_3; \theta)}{\partial \theta^2} = \frac{-(n_1 + n_2)}{\theta^2} - \frac{4n_3}{(1-2\theta)^2}$$
   d'où
$$E\left[\frac{\partial^2 \ln L(N_1, N_2, N_3; \theta)}{\partial \theta^2}\right] = \frac{-2n\theta}{\theta^2} - \frac{4n(1-2\theta)}{(1-2\theta)^2} = \frac{-2n}{\theta(1-2\theta)}.$$
   On en déduit que la borne de Cramér-Rao pour un estimateur non-biaisé de θ est
$$\text{BCR} = \frac{-1}{E\left[\frac{\partial^2 \ln L(N_1,N_2,N_3;\theta}{\partial\theta^2}\right]} = \frac{\theta(1-2\theta)}{2n}.$$
   Comme 1) $\widehat{\theta}_{\text{MV}}$ est un estimateur non biaisé de θ et 2) $\text{var}\left[\widehat{\theta}_{\text{MV}}\right] = \text{BCR}$, on en déduit que $\widehat{\theta}_{\text{MV}}$ est l'estimateur efficace du paramètre θ.

4. (5pts) On suppose maintenant qu'on dispose d'une information a priori sur le paramètre θ résumée dans la loi a priori $p(\theta)$ définie par une loi Beta à trois paramètres notée $\mathcal{B}(\alpha, \beta, \gamma)$ de densité
$$\pi(\theta) = \frac{\theta^{\alpha-1}(\gamma - \theta)^{\beta-1}}{\gamma^{\alpha+\beta-1}\text{B}(\alpha,\beta)}\mathcal{I}_{]0,\gamma[}(\theta)$$
   avec $\gamma = \frac{1}{2}$, $\text{B}(\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$ et où $\mathcal{I}_{]0,\gamma[}$ est la fonction indicatrice sur $]0,\gamma[$ ($\mathcal{I}_{]0,\gamma[}(x) = 1$ si $x \in]0,\gamma[$ et $\mathcal{I}_{]0,\gamma[}(x) = 0$ sinon). On admettra qu'une telle loi est de moyenne $\frac{\alpha\gamma}{\alpha+\beta}$ et de variance $\frac{\gamma^2\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta-1)}$.

   - La loi a priori de θ est de support $]0, 1/2[$. Pourquoi ne peut-on pas avoir $\theta > 1/2$?
   - Déterminer l'estimateur MAP de θ obtenu à l'aide de la vraisemblance $L(n_1, n_2, n_3; \theta)$ et de la loi a priori $\pi(\theta)$ noté $\widehat{\theta}_{\text{MAP}}$.
   - Expliquer le comportement de $\widehat{\theta}_{\text{MAP}}$ lorsque n_1, n_2 et n_3 tendent vers $+\infty$.
   - Montrer que la loi a posteriori de $\theta|n_1, n_2, n_3$ est une loi beta à trois paramètres (on déterminera ces trois paramètres). En déduire l'estimateur de la moyenne a posteriori du paramètre θ noté $\widehat{\theta}_{\text{MMSE}}$.

   - On a $p_3 = 1 - 2\theta \geq 0$ donc $\theta \leq \frac{1}{2}$.

- La loi a posteriori du paramètre θ vérifie

$$p(\theta|n_1,n_2,n_3) \propto p(n_1,n_2,n_3|\theta)\pi(\theta)$$
$$\propto \theta^{n_1+n_2}(1-2\theta)^{n_3}\theta^{\alpha-1}\left(\frac{1}{2}-\theta\right)^{\beta-1}\mathcal{I}_{]0,\frac{1}{2}[}(\theta),$$

soit

$$p(\theta|n_1,n_2,n_3) \propto \theta^{n_1+n_2+\alpha-1}\left(\frac{1}{2}-\theta\right)^{n_3+\beta-1}\mathcal{I}_{]0,\frac{1}{2}[}(\theta).$$

La log-posterior admet donc comme dérivée

$$\frac{\partial \ln p(\theta|n_1,n_2,n_3)}{\partial\theta} = \frac{n_1+n_2+\alpha-1}{\theta} - \frac{n_3+\beta-1}{\frac{1}{2}-\theta}$$

qui s'annule pour

$$\theta = \frac{n_1+n_2+\alpha-1}{2(n+\alpha+\beta-2)}.$$

On en déduit que l'estimateur du maximum a posteriori du paramètre θ est

$$\widehat{\theta}_{\text{MAP}} = \frac{N_1+N_2+\alpha-1}{2(n+\alpha+\beta-2)}.$$

- On remarque que l'estimateur du maximum a posteriori du paramètre θ s'écrit

$$\widehat{\theta}_{\text{MAP}} = \widehat{\theta}_{\text{MV}}\frac{1+\frac{\alpha-1}{N_1+N_2}}{1+\frac{\alpha+\beta-2}{n}}$$

avec $n = N_1+N_2+N_3$. On peut donc remarquer que $\widehat{\theta}_{\text{MAP}}$ et $\widehat{\theta}_{\text{MV}}$ se comportent de manière similaire lorsque $N_1,N_2,N_3 \to +\infty$. Lorsqu'on a beaucoup de données, on a tendance à oublier l'information a priori et à faire confiance à l'estimateur du maximum de vraisemblance qui résume l'information contenue dans les données.

- On a déjà obtenu

$$p(\theta|n_1,n_2,n_3) \propto \theta^{n_1+n_2+\alpha-1}\left(\frac{1}{2}-\theta\right)^{n_3+\beta-1}\mathcal{I}_{]0,\frac{1}{2}[}(\theta).$$

qui est une loi Beta à trois paramètres $\mathcal{B}\left(n_1+n_2+\alpha, n_3+\beta, \gamma=\frac{1}{2}\right)$. D'après les rappels sur cette loi, on en déduit

$$\widehat{\theta}_{\text{MMSE}} = E\left[\theta|N_1,N_2,N_3\right] = \frac{N_1+N_2+\alpha}{2(n+\alpha+\beta)}.$$

Exercice 2 : Test Statistique (10 points)

On considère un triplet de variables aléatoires (N_1, N_2, N_3) de loi multinomiale définie par $P[N_1 = n_1, N_2 = n_2, N_3 = n_3] = \frac{n!}{n_1!n_2!n_3!}p_1^{n_1}p_2^{n_2}p_3^{n_3}$ avec $n_1+n_2+n_3 = n$ et $\boldsymbol{p} = (p_1, p_2, p_3) = (\theta, \theta, 1-2\theta)$. On admettra que pour un tel modèle statistique, la loi marginale de N_i est une loi binomiale $B(n, p_i)$ de moyenne $E[N_i] = np_i$ et de variance $\text{var}(N_i) = np_i(1 - p_i)$. On veut à l'aide de l'observation de (N_1, N_2, N_3) tester les deux hypothèses suivantes

$$H_0 : \theta = \theta_0 \text{ contre } H_1 : \theta = \theta_1 \qquad \text{avec } \theta_1 > \theta_0.$$

1. (2pts) À l'aide du théorème de Neyman-Pearson, montrer que la statistique du test le plus puissant est $T_n = N_3$ (on n'oubliera pas que $N_1 + N_2 + N_3 = n$)) et indiquer la région critique de ce test.
   Le test de Neyman Pearson est défini par

$$\text{Rejet de } H_0 \text{ si } \frac{L(n_1, n_2, n_3; \theta_1)}{L(n_1, n_2, n_3; \theta_0)} > S_{1,\alpha}$$

   où $S_{1,\alpha}$ est un seuil dépendant du risque de première espèce α. Mais

$$\begin{aligned} \frac{L(n_1, n_2, n_3; \theta_1)}{L(n_1, n_2, n_3; \theta_0)} > S_{1,\alpha} &\Leftrightarrow \ln\left[\frac{L(n_1, n_2, n_3; \theta_1)}{L(n_1, n_2, n_3; \theta_0)}\right] > S_{2,\alpha} \\ &\Leftrightarrow (n_1 + n_2) \ln\left(\frac{\theta_1}{\theta_0}\right) + n_3 \ln\left(\frac{1 - 2\theta_1}{1 - 2\theta_0}\right) > S_{2,\alpha}. \end{aligned}$$

   En utilisant $n_1 + n_2 = n - n_3$, on obtient

$$\text{Rejet de } H_0 \text{ si } n_3 \ln\left(\frac{\theta_0(1 - 2\theta_1)}{\theta_1(1 - 2\theta_0)}\right) > S_{3,\alpha}.$$

   Comme $\theta_1 > \theta_0$, on a $\frac{\theta_0(1-2\theta_1)}{\theta_1(1-2\theta_0)} < 1$ donc $\ln\left(\frac{\theta_0(1-2\theta_1)}{\theta_1(1-2\theta_0)}\right) < 0$, d'où la règle de décision

$$\text{Rejet de } H_0 \text{ si } N_3 < S_\alpha.$$

   La région critique du test est l'ensemble des vecteurs $(n_1, n_2, n_3) \in \mathbb{N}^3$ tels que $n_3 < S_\alpha$ et la statistique de test est $T_n = N_3$.

2. (1pt) En utilisant le théorème central limite, déterminer la loi asymptotique de T_n sous les deux hypothèses.
   La loi de T_n est une loi binomiale de moyenne $n(1 - 2\theta)$ et de variance $2n\theta(1 - 2\theta)$. On sait que T_n peut s'écrire comme une somme de variables indépendantes de lois de Bernoulli. L'application du théorème central limite permet alors d'obtenir les résultats suivants

$$\text{Sous } H_0 : T_n \approx \mathcal{N}\left(n(1 - 2\theta_0), 2n\theta_0(1 - 2\theta_0)\right),$$

$$\text{Sous } H_1 : T_n \approx \mathcal{N}\left(n(1 - 2\theta_1), 2n\theta_1(1 - 2\theta_1)\right),$$

   où $\approx$ signifie "de loi approchée (pour n grand)".

3. (2pts) En utilisant la loi asymptotique de la question précédente, déterminer les risques de première et seconde espèce α et β du test en fonction des paramètres θ_0 et θ_1 et de la fonction de répartition d'une loi normale $\mathcal{N}(0, 1)$ notée F.
   Le risque α est défini par

$$\alpha = P[\text{Rejeter } H_0 | H_0 \text{ vraie}] = P\left[T_n < S_\alpha | T_n \sim \mathcal{N}\left(n(1 - 2\theta_0), 2n\theta_0(1 - 2\theta_0)\right)\right],$$

   soit

$$\alpha = F\left[\frac{S_\alpha - n(1 - 2\theta_0)}{\sqrt{2n\theta_0(1 - 2\theta_0)}}\right].$$

De même

$$\beta = P[\text{Rejeter } H_1 | H_1 \text{ vraie}] = P\left[T_n \geq S_\alpha | T_n \sim \mathcal{N}\left(n(1-2\theta_1), 2n\theta_1(1-2\theta_1)\right)\right],$$

soit

$$\beta = 1 - F\left[\frac{S_\alpha - n(1-2\theta_1)}{\sqrt{2n\theta_1(1-2\theta_1)}}\right].$$

4. (2pts) Déterminer les courbes COR associées à ce test et tracer l'allure de ces courbes pour différentes valeurs de n.
   Les courbes COR expriment $\pi = 1 - \beta$ en fonction de α. On a donc

$$\pi = F\left[\frac{S_\alpha - n(1-2\theta_1)}{\sqrt{2n\theta_1(1-2\theta_1)}}\right]$$

   avec $S_\alpha = F^{-1}(\alpha)\sqrt{2n\theta_0(1-2\theta_0)} + n(1-2\theta_0)$, soit

$$\pi = F\left[F^{-1}(\alpha)\sqrt{\frac{\theta_0(1-2\theta_0)}{\theta_1(1-2\theta_1)}} + \sqrt{2n}\frac{\theta_1 - \theta_0}{\sqrt{\theta_1(1-2\theta_1)}}\right].$$

   On observe donc que π est une fonction croissante de n, ce qui est habituel car plus le nombre d'observations est grand, meilleure est la performance du test. L'allure des courbes COR pour différentes valeurs de n est représentée sur la figure ci-dessous

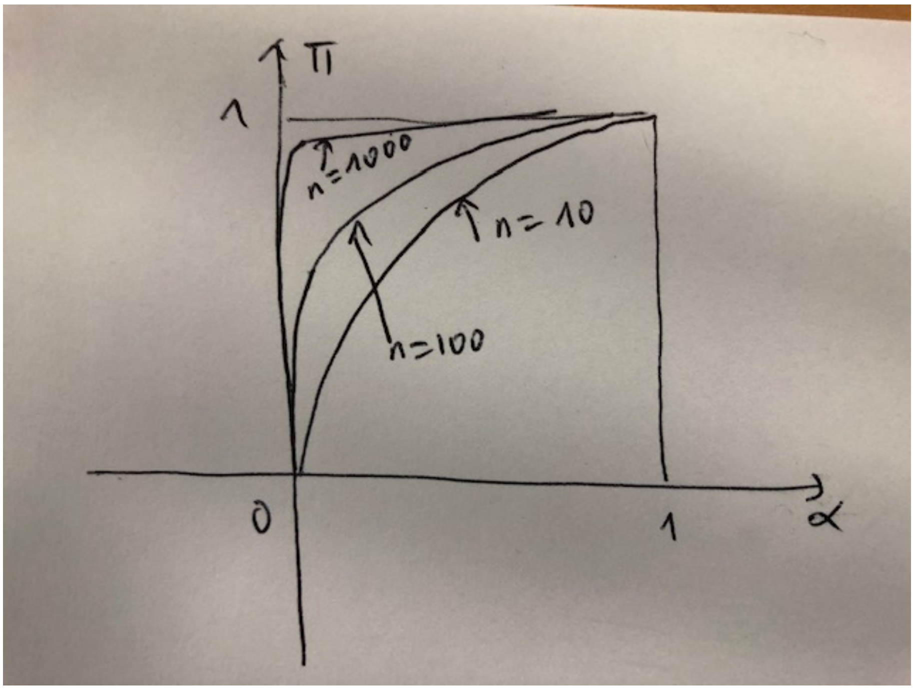

Figure 1: Allure des courbes COR pour différentes valeurs de n.

5. (3pts) On désire à partir d'un ensemble de $K = 10$ observations de N_1 notées $(n_{11}, ..., n_{1K})$ et données dans le tableau ci-dessous déterminer s'il est raisonnable de supposer que N_1 suit une loi binomiale $B(n, \theta)$ avec $n = 10$.

2	3	3	2	2	0	4	3	1	0

- Expliquer pourquoi on est amené à choisir $\theta = 0.2$ pour effectuer ce test.
  Comme θ est inconnu, il faut l'estimer avec l'estimateur du maximum de vraisemblance construit à partir des données $n_{11}, ..., n_{1K}$, soit

$$\hat{\theta} = \frac{1}{nK}\sum_{k=1}^{K} n_{1k} = \frac{20}{10 \times 10} = 0.2.$$

- Les probabilités associées à une loi binomiale $B(n, \theta)$ avec $\theta = 0.2$ et $n = 10$ sont données dans le tableau ci-dessous

p_0	p_1	p_2	p_3	p_4	$\sum_{k=5}^{10} p_k$
0.11	0.27	0.30	0.20	0.09	0.03

  On décide de faire un test du χ^2 avec 3 classes notées C_1, C_2 et C_3. Justifier le choix suivant pour ces trois classes : $C_1 = \{0, 1\}$, $C_2 = \{2\}$ et $C_1 = \{3, ..., 10\}$.
  Comme la loi binomiale est discrète, il faut regrouper quelques valeurs de $\{0, ..., n\}$ dans chaque classe, de manière à avoir des classes les plus équiprobables possibles. On en déduit les trois classes $C_1 = \{0, 1\}$, $C_2 = \{2\}$ et $C_1 = \{3, ..., 10\}$ avec les probabilités $P_1 = p_1 + p_2 = 0.38$, $P_2 = 0.3$ et $P_3 = 0.32$. On remarquera que la condition $np_k \geq 5$ n'est pas vérifiée et donc on devrait surement choisir un autre test !.

- Exprimer la statistique de ce test notée ϕ en fonction des données du problème (sans chercher à la calculer) et préciser la région critique.
  La statistique de test est classiquement définie par

$$\phi = \sum_{c=1}^{C} \frac{(z_c - KP_c)^2}{KP_c},$$

  où $C = 3$ est le nombre de classes, P_c est la somme des probabilités des éléments appartenant à la classe $\#c$, K est le nombre d'observations de la variable N_1 et z_c est le nombre de ces observations appartenant à la classe $\#c$. On obtient donc

$$\phi = \frac{(3 - 3.8)^2}{3.8} + \frac{(3 - 3)^2}{3} + \frac{(4 - 3.2)^2}{3.2}.$$

  La région critique est définie par : Rejet de H_0 si $\phi > S_\alpha$, où α est le risque de première espèce et S_α le seuil associé.

- Quelle est la loi asymptotique de ϕ ?
  La loi asymptotique de ϕ est une loi du χ^2_{C-2} car on doit estimer le paramètre $p_1 = \theta$ à l'aide des données n_{1k}, c'est-à-dire $\phi \sim \chi^2_1$.

- Les seuils obtenus pour $\alpha = 0.05$ et $\alpha = 0.01$ sont $S_{0.05} = 5.99$ et $S_{0.01} = 9.21$. Expliquer pourquoi $S_{0.05} < S_{0.01}$.
  Lorsque α diminue, la probabilité de rejeter H_0 diminue, donc le seuil S_α augmente.

LOIS DE PROBABILITÉ CONTINUES

$\mathbf{m}$: moyenne $\quad$ σ^2 : variance $\quad$ **F. C. :** fonction caractéristique

LOI	**Densité de probabilité**	**m**	σ^2	**F. C.**
Uniforme	$f(x) = \frac{1}{b-a}$ $x \in \,]a, b[$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\dfrac{e^{itb} - e^{ita}}{it(b-a)}$
Gamma $\mathcal{G}(\nu, \theta)$	$f(x) = \frac{\theta^\nu}{\Gamma(\nu)} e^{-\theta x} x^{\nu-1}$ $\theta > 0,\ \nu > 0$ $x \geq 0$ avec $\Gamma(n+1) = n!\ \forall n \in \mathbb{N}$	$\dfrac{\nu}{\theta}$	$\dfrac{\nu}{\theta^2}$	$\dfrac{1}{\left(1 - i\frac{t}{\theta}\right)^\nu}$
Inverse gamma $\mathcal{IG}(\nu, \theta)$	$f(x) = \frac{\theta^\nu}{\Gamma(\nu)} e^{-\frac{\theta}{x}} \frac{1}{x^{\nu+1}}$ $\theta > 0,\ \nu > 0$ $x \geq 0$ avec $\Gamma(n+1) = n!\ \forall n \in \mathbb{N}$	$\frac{\theta}{\nu-1}$ si $\nu > 1$	$\frac{\theta^2}{(\nu-1)^2(\nu-2)}$ si $\nu > 2$	(*)
Première loi de Laplace	$f(x) = \frac{1}{2} e^{-\lvert x \rvert}, \quad x \in \mathbb{R}$	0	2	$\dfrac{1}{1+t^2}$
Normale univariée $\mathcal{N}\left(m, \sigma^2\right)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad x \in \mathbb{R}$	m	σ^2	$e^{imt - \frac{\sigma^2 t^2}{2}}$
Normale multivariée $\mathcal{N}_p(\boldsymbol{m}, \boldsymbol{\Sigma})$	$f(x) = K \exp\left[-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{m})^T \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{m})\right]$ $K = \frac{1}{\sqrt{(2\pi)^p \det(\Sigma)}}$ $x \in \mathbb{R}^p$	$\boldsymbol{m}$	$\boldsymbol{\Sigma}$	$e^{i\boldsymbol{u}^T\boldsymbol{m} - \frac{1}{2}\boldsymbol{u}^T\boldsymbol{\Sigma}\boldsymbol{u}}$
Khi$_2$ χ^2_ν $\Gamma\left(\frac{1}{2}, \frac{\nu}{2}\right)$	$f(x) = k e^{-\frac{x}{2}} x^{\frac{\nu}{2}-1}$ $k = \dfrac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)}$ $\nu \in \mathbb{N}^*,\ x \geq 0$	ν	2ν	$\dfrac{1}{(1-2it)^{\frac{\nu}{2}}}$
Cauchy $c_{\lambda,\alpha}$	$f(x) = \dfrac{1}{\pi\lambda\left(1 + \left(\frac{x-\alpha}{\lambda}\right)^2\right)}$ $\lambda > 0,\ \alpha \in \mathbb{R}$	(-)	(-)	$e^{i\alpha t - \lambda\lvert t \rvert}$
Beta $B(a, b)$	$f(x) = k x^{a-1}(1-x)^{b-1}$ $k = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$ $a > 0,\ b > 0$ $x \in \,]0, 1[$ avec $\Gamma(n+1) = n!\ \forall n \in \mathbb{N}$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$	(*)

LOIS DE PROBABILITÉ DISCRÈTES
$\mathbf{m}$: moyenne σ^2 : variance **F. C. :** fonction caractéristique
$p_k = P[X = k]$ $p_{1,\ldots,m} = P[X_1 = k_1, \ldots, X_m = k_m]$

LOI	Probabilités	$\mathbf{m}$	σ^2	F. C.
Uniforme	$p_k = \frac{1}{n}$ $k \in \{1, \ldots, n\}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$	$\dfrac{e^{it}\left(1 - e^{itn}\right)}{n\left(1 - e^{it}\right)}$
Bernoulli	$p_1 = P[X = 1] = p$ $p_0 = P[X = 0] = q$ $p \in [0, 1] \quad q = 1 - p$	p	pq	$pe^{it} + q$
Binomiale $B(n, p)$	$p_k = C_n^k p^k q^{n-k}$ $p \in [0, 1] \quad q = 1 - p$ $k \in \{0, 1, \ldots, n\}$	np	npq	$\left(pe^{it} + q\right)^n$
Binomiale négative	$p_k = C_{n+k-1}^{n-1} p^n q^k$ $p \in [0, 1] \quad q = 1 - p$ $k \in \mathbb{N}$	$n\dfrac{q}{p}$	$n\dfrac{q}{p^2}$	$\left(\dfrac{p}{1 - qe^{it}}\right)^n$
Multinomiale	$p_{1,\ldots,m} = \frac{n!}{k_1! \ldots k_m!} p_1^{k_1} \ldots p_m^{k_m}$ $p_j \in [0, 1] \quad q_j = 1 - p_j$ $k_j \in \{0, 1, \ldots, n\}$ $\sum_{j=1}^m k_j = n \quad \sum_{j=1}^m p_j = 1$	np_j	Variance : $np_j q_j$ Covariance : $-np_j p_k$	$\left(\sum_{j=1}^m p_j e^{it}\right)^n$
Poisson $P(\lambda)$	$p_k = e^{-\lambda}\dfrac{\lambda^k}{k!}$ $\lambda > 0 \quad k \in \mathbb{N}$	λ	λ	$\exp\left[\lambda\left(e^{it} - 1\right)\right]$
Géométrique	$p_k = pq^{k-1}$ $p \in [0, 1] \quad q = 1 - p$ $k \in \mathbb{N}^*$	$\dfrac{1}{p}$	$\dfrac{q}{p^2}$	$\dfrac{pe^{it}}{1 - qe^{it}}$