Axiomes de la théorie des probabilités

Quels que soient les évènements $\mathcal{E}$ et $\mathcal{E}'$ de l'espace de tous les résultats possibles Ω:

- $0 \leq P(\mathcal{E}) \leq 1$ et $P(\Omega) = 1$
- $\mathcal{E} \cap \mathcal{E}' = \emptyset \Rightarrow P(\mathcal{E} \cup \mathcal{E}') = P(\mathcal{E}) + P(\mathcal{E}')$
- $\mathcal{E} \subset \mathcal{E}' \Rightarrow P(\mathcal{E}) \leq P(\mathcal{E}')$

Variance, covariance, linéarité de l'espérance mathématique

- $\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$
- σ_X : $\mathrm{std}(X) = \sqrt{\mathrm{Var}(X)}$
- $\mathrm{cov}(X,Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$
- $\mathrm{corr}(X,Y) = \mathrm{cov}(X,Y)/(\sigma_X\, \sigma_Y)$
- $[ax + b] = a[x] + b$
- $\mathrm{Var}(ax + b) = a^2\, \mathrm{Var}(x)$
- $\mathrm{std}(ax + b) = |a|\, \mathrm{std}(x)$
- $\mathrm{cov}(ax + b, cy + d) = a\, x\, \mathrm{cov}(x, y)$
- $\mathrm{corr}(ax + b, cy + d) = \mathrm{corr}(x, y)$
- x, y indép. $\Rightarrow \mathrm{cov}(x, y) = 0$
- $\mathrm{Var}(x + y) = \mathrm{Var}(x) + 2\mathrm{cov}(x, y) + \mathrm{Var}(y)$
- x, y indép. $\Rightarrow \mathrm{Var}(x + y) = \mathrm{Var}(x) + \mathrm{Var}(y)$
- $\forall x,\ \forall y,\ |\mathrm{corr}(x, y)| \leq 1$

Si X prend des valeurs dans $\mathbb{R}^d$ (cas multidimensionnel), la matrice de covariance s'écrit

$$\mathrm{Cov}(X) = \mathbb{E}[(X - \mathbb{E}(X))(X - \mathbb{E}(X))^T]$$

Fonction de répartition et densité de probabilité

F_X est la fonction de répartition.
p_X est la densité de probabilité ou la distribution de probabilité.

- $F_X(x) \triangleq P(X \leq x)$
- p_X satisfait $F_X(x) = \int_{-\infty}^{x} p_X(\xi) d\xi$
- Si F_x est dérivable au point x (la dérivée est définie en x) alors $\frac{dF_x}{dx} = p_X(x)$

Espérance mathématique d'une fonction d'une variable aléatoire

$$\mathbb{E}[h(X)] = \int h(\xi)\, p_X(\xi)\, d\xi$$

Quantiles d'une variable aléatoire continue

Si F_X est inversible (i.e. F_X^{-1} existe) alore le quantile de niveau α est défini par

$$q_\alpha \triangleq F_X^{-1}(\alpha)$$

Distribution jointe d'une paire de variables aléatoires continues

- Fonction de répartition jointe $F(x, y) \triangleq P(X \leq x, Y \leq y)$
- Densité de probabilité jointe: f tel que

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(\xi, \eta) d\xi\, d\eta$$

- Densité de probabilité conditionnelle: $p_{Y|X}(y|x) \triangleq \frac{p_{X,Y}(x,y)}{p_X(x)}$
- Loi de la probabilité totale: $p_X(x) = \int p_{X,Y}(x, y) dy$
- Règle de Bayes : $p_{Y|X}(y|x) = \frac{p_{X|Y}(x|y) p_Y(y)}{p_X(x)}$
- Espérance conditionnelle: $\mathbb{E}[X|Y = y] = \int x p_{X|Y}(x|y) dx$
- $\mathbb{E}[X|Y] = \int x p_{X|Y}(x|Y) dx$
- Loi de l'espérance mathématique totale: $\mathbb{E}[[E|Y]] = \mathbb{E}[X]$

Indépendance

Les variables aléatoires continues X et Y sont indépendantes si et seulement si n'importe laquelle des conditions énumérées suivantes est satisfaite:

1. $\forall (x, y) \in \mathcal{X} \times \mathcal{Y}, p_{X,Y}(x, y) = p_X(x)\, p_Y(y)$
2. $\forall (x, y) \in \mathcal{X} \times \mathcal{Y}$, tels que $p_Y(y) > 0$, $p_{X|Y}(x|y) = p_X(x)$
3. $\forall f, g$, $\mathbb{E}[f(X)\, g(X)] = \mathbb{E}[f(X)]\, \mathbb{E}[g(X)]$

Une collectin de variables aléatoires X_1, ..., X_n sont indépendantes, si, et seulement si, une des trois conditions énumérées suivantes est satisfaite (notation: $X_{-i} \triangleq (X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$)

1. $\forall x \in \mathbb{R}^n, p_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = p_{X_1}(x_1) \cdots p_{X_n}(x_n)$
2. $\forall x \in \mathbb{R}^n$ tels que $p_{X_{-i}}(x_{-i}) > 0$, $\forall i$, $p_{X_i|X_{-i}}(x_i|x_{-i}) = p_{X_i}(x_i)$
3. $\forall f_1, \ldots, f_n$, $\mathbb{E}[f(X_1) \cdots f_n(X_n)] = \prod_{i=1}^n \mathbb{E}[f_i(X)]$

Les mêmes propriétés sont satisfaites pour les variables aléatoires discrètes une fois que l'on remplace les densités de probabilités continues par les distributions de probabilité discrète.

Somme de variables aléatoires et convolution des densités

Si $Z = X + Y$ avec X, Y des variables alétoires continues (resp. discrètes) *indépendantes*

$$p_Z(z) = (p_X * p_Y)(z) \triangleq \int_{-\infty}^{+\infty} p_X(z - y) p_Y(y) dy$$

$$P_Z(z) = (P_X * P_Y)(z) \triangleq \sum_{y=-\infty}^{+\infty} P_X(x - y)\, P_Y(y)$$

Décalage, mise à l'échelle et transformation de variables aléatoires

- Si $Y = aX + b$ alors $p_Y(y) = \frac{1}{a} p_X\left(\frac{Y-b}{a}\right)$
- Si $Y = f(X)$ avec f strictement monotone alors

$$p_Y(y) = p_X(f^{-1}(y))\,|(f^{-1})'(y)|$$

- Si $Z \sim \mathcal{N}(0,1)$ alors $X^2 \sim \chi^2_{(1)}$

Somme de variables aléatoires indépendantes et identiquement distribuées (i.i.d.)

Si $X_i \overset{\text{iid}}{\sim}$ ...	alors $\sum_i X_i \sim$ ...
Ber(p)	Bin(n,p)
Pois(θ)	Pois$(n\,\theta)$
$\mathcal{N}(\mu,\sigma^2)$	$\mathcal{N}(n\,\mu, n\,\sigma^2)$
Exponentielle $\mathcal{E}(\lambda)$	$\Gamma(n,\lambda)$
$\Gamma(r,\lambda)$	$\Gamma(nr,\lambda)$
$\chi^2_{(1)}$	$\chi^2_{(n)}$

Pour $X_i \sim \Gamma(r_i,\lambda)$ indép., $\sum_{i=1}^n X_i \sim \Gamma(\sum_{i=1}^n r_i, \lambda)$
Stabilité de la famille Gaussienne: Si X_1 et X_2 sont indép. et $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, alors $a_1X_1 + a_2X_2 \sim \mathcal{N}(a_1\mu_1 + a_2\mu_2, a_1^2\sigma_1^2 + a_2^2\sigma_2^2)$

Fonctions génératrices des moments

$$M_X(u) = \mathbb{E}[e^{uX}] = \int_{-\infty}^{+\infty} e^{ux}\, f(x)\, dx$$

où $u \in \mathbb{R}$ et f est la densité de probabilité (la distribution de probabilité).

propriété	formule
linéarité	$a_1\, f_1(x) + a_2\, f_2(x) \leftrightarrow a_1 M_1(u) + a_1 M_2(u)$
décalage en x	$f(x - x_0) \leftrightarrow e^{u x_0}\, M(u)$
décalage en u	$e^{-u_0 x}\, f(x) \leftrightarrow M(u - u_0)$
mise en échelle en x	$f(a\, x) \leftrightarrow \frac{1}{\lvert a \rvert} M(u)$
renversement	$f(-x) \leftrightarrow M(-u)$
dérivée par rapport à x	$\frac{df(x)}{dx} \leftrightarrow -uM(u)$
dérivée par rapport à u	$x\, f(x) \leftrightarrow \frac{dM(u)}{du}$
intégration	$\int_{-\infty}^{t} f(\xi)\, d\xi \leftrightarrow -\frac{1}{u} M(u)$
convolution	$f_1(x) * f_2(x) \leftrightarrow M_1(u) \cdot M_2(u)$

Fonctions caractéristiques

$$
\begin{aligned}
f_X(x) &\overset{\mathcal{F}}{\leftrightarrow} \phi_X(u) \\
\phi_X(u) &\triangleq \mathbb{E}[e^{+jux}] = \int_{-\infty}^{+\infty} f_X(x)\, e^{jux}\, dx \\
f_X(x) &\triangleq \frac{1}{2\pi} \int_{-\infty}^{+\infty} \phi(u)\, e^{-jux}\, du
\end{aligned}
$$

propriété	correspondance
linéarité	$a_1 f_1(x) + a_2 f_2(x) \leftrightarrow a_1\, \phi_1(u) + a_2\, \phi_2(u)$
décalage en x	$f(x - x_0) \leftrightarrow e^{jux_0}\, \phi(u)$
décalage en u	$e^{-ju_0 x}\, f(x) \leftrightarrow \phi(u - u_0)$
mise en échelle en x	$f(a\, x) \leftrightarrow \frac{1}{\lvert a \rvert} \phi\left(\frac{u}{a}\right)$
renversement	$f(-x) \leftrightarrow \phi(-u)$
dualité	$\phi(t) \leftrightarrow 2\pi\, f(u)$
dérivée selon x	$\frac{df}{dx} \leftrightarrow -ju\, \phi(u)$
dérivée selon u	$j\, x\, f(x) \leftrightarrow \frac{d\phi(u)}{du}$
intégration	$\int_{-\infty}^{x} f(\xi)\, d\xi \leftrightarrow \pi\phi(0)\, \delta(u) - \frac{1}{ju}\phi(u)$
convolution	$f_1 * f_2 \leftrightarrow \phi_1(u) \cdot \phi_2(u)$
multiplication	$f_1(x) \cdot f_2(x) \leftrightarrow \frac{1}{2\pi}\, \phi_1(u) * \phi_2(u)$
Parseval	$\int_{-\infty}^{+\infty} f_1(x)\, f_2(x)\, dx = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X_1(u)\, X_2(-u)\, du$

Tableau récapitulatif

dénomination	fonction de probabilité	fct. gén. des moments $M_X(u)$	fct. caractéristique $\phi(u)$
Ponctuelle (point massique en a)	$\delta(x-a)$	e^{ua}	e^{iua}
Bernouilli	$P(x=1)=p$	$1-p+p\cdot e^u$	$1-p+p\,e^{iu}$
Binomiale $\mathcal{B}(n,p)$	$\begin{pmatrix} n \\ x \end{pmatrix} p^x(1-p)^{n-x}$	$(1-p+pe^u)^n$	$(1-p+pe^{iu})^n$
Géométrique	$(1-p)^k p$	$\frac{p}{1-(1-p)e^u}$ $u<-\ln(1-p)$	$\frac{p}{1-(1-p)}e^{iu}$
Poisson $\mathcal{P}oi(\lambda)$	$\frac{\lambda^k}{k!}e^{-\lambda}$	$e^{\lambda(e^u-1)}$	$e^{\lambda(e^{iu}-1)}$
Uniforme $U(a,b)$	$\frac{1}{b-a}\left[\epsilon(x-a)-\epsilon(x-b)\right]$	$\frac{e^{ub}-e^{ua}}{u(b-a)}$	$\frac{e^{iub}-e^{iua}}{iu(b-a)}$
Normale $\mathcal{N}(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$e^{u\mu+\sigma^2\frac{u^2}{2}}$	$e^{iu\mu-\sigma^2\frac{u^2}{2}}$
$\chi^2_{(\nu)}$	$\frac{1}{2^{\nu/2}\,\Gamma(\nu/2)}\cdot x^{\frac{\nu}{2}-1}\cdot e^{-\frac{x}{2}}$	$(1-2u)^{-\nu/2}$, $u<1/2$	$(1-2iu)^{-\nu/2}$
$\Gamma(k,1/\theta)$	$\frac{1}{\Gamma(\alpha)}(1/\theta)^\alpha x^{\alpha-1}\,e^{-x/\theta}$	$(1-\theta\,u)^{-k}$, $u<1/\theta$	$(1-i\theta\,u)^{-k}$
$\mathrm{Exp}(\lambda)$	$\lambda\,e^{-\lambda\,x}$	$(1-u\lambda^{-1})^{-1}$, $u<\lambda$	$(1-ui\lambda^{-1})^{-1}$

Estimation

- $\text{Biais}(\hat{\theta}, \theta) \triangleq \mathbb{E}[\hat{\theta}] - \theta$ et $\text{Var}(\hat{\theta}) \triangleq \mathbb{E}[(\hat{\theta} - \mathbb{E}[\hat{\theta}])^2]$
- Erreur quadratique moyenne: $\text{MSE}(\hat{\theta}, \theta) \triangleq \mathbf{E}[(\hat{\theta} - \theta)^2] = \text{Var}(\hat{\theta}) + \text{Bias}(\hat{\theta}, \theta)^2$
- Méthode du maximum de vraisemblance (MLE): $\hat{\theta}_{\text{MLE}}$

$$\hat{\theta}_{\text{MLE}} = \arg\max_{\theta} l(\theta)$$

avec $l(\theta) \triangleq \sum_{i=1}^{n} \log p(x_i, \theta)$

MLE pour le modèle gaussien

- $\hat{\mu}_{\text{MLE}} = \bar{x}$
- lorsque μ est connu, $\hat{\sigma}_{\text{MLE}} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2$
- lorsque μ est inconnu, $\hat{\theta}_{\text{MLE}} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 = \overline{x^2} - \bar{x}^2$
- variance d'échantillon: $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$

Lorsque les données sont distribuées de manière normale (gaussienne), alors :

- $\hat{\mu}_{\text{MLE}} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$
- lorsque μ est connu, $n\,\hat{\sigma}^2_{\text{MLE}}/\sigma^2 \sim \chi^2_{(n)}$
- lorsque μ est inconnu, $n\,\hat{\sigma}^2_{\text{MLE}}/\sigma^2 = (n-1)s^2/\sigma^2 \sim \chi^2_{(n-1)}$

Intervalles de confiance

Pour $|z_\alpha| = z_{1-\alpha}$ le quantile de niveau $1-\alpha$ d'une distribution normale $\mathcal{N}(0,1)$ et $\tau_\alpha \triangleq t_{(1-\alpha),(n-1)}$, le quantile de niveau $1-\alpha$ d'une distribution de Student à $n-1$ degrés de liberté, les intervalles de confiance ont la forme suivante:

- Gaussienne:
$$\left[\bar{X} - |z_{\alpha/2}|\frac{\sigma}{\sqrt{n}}\,;\,\bar{X} + |z_{\alpha/2}|\frac{\sigma}{\sqrt{n}}\right]$$
- Student:
$$\left[\bar{X} - \tau_{\alpha/2}\frac{S}{\sqrt{n}}\,;\,\bar{X} + \tau_{\alpha/2}\frac{S}{\sqrt{n}}\right]$$
- Asymptotique:
$$\left[\bar{X} - |z_{\alpha/2}|\frac{S}{\sqrt{n}}\,;\,\bar{X} + |z_{\alpha/2}|\frac{S}{\sqrt{n}}\right]$$

Tests d'hypothèses selon Neyman-Pearson

H_0: hypothèse nulle est une présuposition tentative concernant un paramètre du modèle. H_1: Une hypothèse alternative concernant ce paramètre.

- $H_0 : \mu \geq \mu_0$ vs $H_1 : \mu < \mu_0$ (test à une queue, la queue inférieur, de gauche)
- $H_0 : \mu \leq \mu_0$ vs $H_1 : \mu > \mu_0$ (test à une queue, la queue supérieure, de droite)
- $H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$ (test à deux queues)

Erreurs de type I et II:

Erreur de type I	rejeter H_0 alors que H_0 est vraie
Erreur de type II	ne pas rejeter H_0 alors que H_0 est fausse

$P(\text{erreur de type I}) = \alpha$
$P(\text{erreur de type II}) = \beta$

Puissance: c'est la probabilité de rejeter H_0 lorsque H_1 est vraie. C'est égal à $1-\beta$.

Statistiques de test pour les tests classiques

- Test gaussien exact pour μ:

$$X_i \sim^{\text{iid}} \mathcal{N}(\mu, \sigma^2) \text{ et } \sigma^2 \text{ connu } T \triangleq \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$$

- Test de Student exact sur μ

$$X_i \sim^{\text{iid}} \mathcal{N}(\mu, \sigma^2) \text{ et } \sigma^2 \text{ inconnu } T \triangleq \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t_{(n-1)}$$

- Test asymptotique gaussien sur $\mu = \mathbb{E}[X_i]$:

$$X_i \sim^{\text{iid}} P, \text{ et } \mathbb{E}[X_i^b] < \infty : T \triangleq \frac{\bar{X} - \mu_0}{\hat{\sigma}/\sqrt{n}} \to \mathcal{N}(0,1)$$

- Test du χ^2 sur un paramètre σ^2 :

$$X_i \sim^{\text{iid}} \mathcal{N}(\mu, \sigma^2) \text{ et } \mu \text{ inconnu } T \triangleq (n-1)\frac{S^2}{\sigma_0^2} \sim \chi^2_{(n-1)}$$

- Paramètre estimé $\hat{\alpha}$ fonction linéaire des valeurs observées y_i, ..., $\hat{y}_n$ (donc gaussiennes). Variance connue:

$$\hat{\alpha} = a_1 y_1 + a_2 y_2 + \ldots + a_n y_n$$

$$\text{Var}(\hat{\alpha}) = \left(\sum a_i^2\right) \text{Var}(Y_i) = \sum a_i^2 \sigma^2$$

$$Z \triangleq \frac{\hat{\alpha} - \alpha}{\sqrt{\sigma^2 \sum a_i^2}} \sim \mathcal{N}(0,1)$$

- Variance inconnue: $s^2 \triangleq \frac{1}{n-q} \sum_{i=1}^n (y_i - \hat{\mu}_i)^2$

$$T \triangleq \frac{\hat{\alpha} - \alpha}{\sqrt{s^2 \sum a_i^2}} \sim t_{(n-q)}$$

Modèle à deux échantillons

- $$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}$$

$$\bar{Y}_1 \sim \mathcal{N}\left(\mu_1, \frac{1}{n_1}\sigma^2\right) \qquad \bar{Y}_2 \sim \mathcal{N}\left(\mu_2, \frac{1}{n_2}\sigma^2\right)$$

$\bar{Y}_1$ et $\bar{Y}_2$ sont indépendants $\Rightarrow$

$$\bar{Y}_2 - \bar{Y}_1 \sim \mathcal{N}\left(\mu_2 - \mu_1, \sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right)$$

$$\boxed{T \triangleq \frac{\hat{\beta} - \beta}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim t_{(n-2)}}$$

Régression linéaire

- inférence sur $\hat{\beta}$

$$\hat{\beta} = \frac{\sigma_{xy}}{\sigma_{xx}} = \frac{\sum(x_i - \bar{x})y_i}{\sigma_{xx}} = \sum a_i y_i$$

avec a_i les constantes

$$a_i = (x_i - \bar{x})/\sigma_{xx} \qquad i = 1, 2, \ldots, n$$

$$\boxed{T \triangleq \frac{\hat{\beta} - \beta}{\sqrt{s^2 \sum a_i^2}} \sim t_{(n-2)}}$$

- inférence sur $\mathbb{E}[y]$ Pour toute valeur de x, la valeur attendue (espérance mathématique) de Y associée est $y = \alpha + \beta x$, avec l'estimateur MLE

$$\hat{y} = \hat{\alpha} + \hat{\beta} x$$

comme $\hat{\alpha} = \bar{y} + \hat{\beta}\bar{x}$

$$\begin{aligned}
\hat{\mu} &= \bar{y} + \hat{\beta}(x - \bar{x}) \\
&= \frac{1}{n}\sum y_i + (x - \bar{x})\sum a_i y_i \\
&= \sum \left[\frac{1}{n} + (x - \bar{x})a_i\right] y_i
\end{aligned}$$

$$\hat{\mu} \sim \mathcal{N}(\mu, \sigma^2 \sum (1/n - (x - \bar{x})^2/\sigma_{xx})$$

$$\boxed{T' \triangleq \frac{\hat{\mu} - \mu}{\sqrt{s^2\left(\frac{1}{n} + \frac{(x-\bar{x})^2}{\sigma_{xx}}\right)}} \sim t_{(n-2)}}$$

$$\boxed{T'' \triangleq \frac{\hat{\alpha} - \alpha}{\sqrt{s^2\left(\frac{1}{n} + \frac{\bar{x}^2}{\sigma_{xx}}\right)}} \sim t_{(n-2)}}$$

Valeurs critiques pour les test classiques

Les valeurs critiques t_c au niveau de signification (NS) α (ou de manière équivalente au niveau de confiance $1-\alpha$) : Tests exacts et asymptotiques avec z_α le quantile de niveau α d'une distribution normale $\mathcal{N}(0,1)$

- queue inférieure (de gauche): $t_c = z_\alpha$
- queue supérieure (de droite): $t_c = z_{1-\alpha} = |z_\alpha|$
- deux queues: $t_c = z_{1-\alpha/2} = |z_{\alpha/2}|$

Pour les tests de Student: avec $t_{(n-1),\alpha}$ le quantile de niveau α de la distribution de Student $t_{(n-1)}$:

- queue inférieure: $t_c = -\tau_\alpha \triangleq t_{(n-1),\alpha}$
- queue supérieure: $t_c = \tau_\alpha \triangleq t_{(n-1),1-\alpha} = |t_{(n-1),\alpha/2}|$

Pour les tests du χ^2, avec $\chi^2_{(n-1),\alpha}$ le quantile de niveau α d'une distribution $\chi^2_{(n-1)}$:

- queue inférieure: $t_c = \chi^2_{(n-1),2}$
- queue supérieure: $t_c = \chi^2_{(n-1),1-\alpha}$
- deux queues: requiert deux seuils, la région critique est

$$\Big]-\infty\,;\,\chi^2_{(n-1),\alpha/2}\Big] \cup \Big[\chi^2_{(n-1),1-\alpha/2}\,;\,+\infty\Big[$$

Tests: règle pour le rejet et région critique $\mathcal{C}$

Soit t_{obs} la valeur de la statistique observée

- queue inférieure: rejeter H_0 lorsque $t_{\text{obs}} \in \mathcal{C} \triangleq]-\infty\,;\,t_c]$
- queue supérieure: rejeter H_0 lorsque $t_{\text{obs}} \in \mathcal{C} \triangleq [t_c\,;\,+\infty[$
- deux queues: rejeter H_0 lorsque $t_{\text{obs}} \in \mathcal{C} \triangleq]-\infty\,;\,-t_c] \cup [t_c\,;\,+\infty[$

Tests: p-valeur p_{obs}

Soit t_{obs} la valeur observée de la statistique, et P_0 la probabilité sous H_0, la p-valeur p_{obs} est

- queue inférieure: $p_{\text{obs}} = P_0(T \leq t_{\text{obs}})$
- queue supérieure: $p_{\text{obs}} = P_0(T \geq t_{\text{obs}})$
- deux queues: $p_{\text{obs}} = P_0(|T| \geq |t_{\text{obs}}|)$

Tests: puissance

Soit P_1 la probabilité sous H_1. La puissance est

- queue inférieure: $1 - \beta = P_1(T \leq t_c)$
- queue supérieure: $1 - \beta = P_1(T \geq t_c)$
- deux queues: $1 - \beta = P_1(|T| \geq t_c)$

Concepts bayésiens et principes

Soit un échantillon $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$

- distribution a priori: $\pi(\theta)$
- vraisemblance: $p(\mathcal{X}|\theta) = \prod_{i=1}^{n} p(x_i|\theta)$
- distribution a posteriori: $\pi(\theta|\mathcal{X} = \frac{p(\mathcal{X}|\theta)\,\pi(\theta)}{p(\mathcal{X})}$ avec

$$p(\mathcal{X}) = \int p(\mathcal{X}|\theta)\,\pi(\theta)\,d\theta$$

- la moyenne a posteriori (posterior mean PM)

$$\theta_{\text{PM}} = \mathbb{E}[\theta|\mathcal{X}] = \int \theta\,\pi(\theta|\mathcal{X})d\theta$$

- le maximum a posteriori

$$\theta_{\text{MAP}} = \arg\max \pi(\theta|\mathcal{X}) = \arg\max_{\theta} l(\theta) + \log \pi(\theta)$$

Variables aléatoires de Bernouilli, binomiale et multinomiale

- Distribution de Bernouilli: $X \sim$ Ber(p)

$$P(X=x) = p^x(1-p)^{1-x} \text{ pour } x = 0.1$$

espérance	$\mathbb{E}[X] = p$
variance	$\mathrm{Var}(X) = p(1-p)$

- Distribution Binomiale: distribution de sommes de n variables aléatoires indépendantes de Bernouilli: $N \sim \mathrm{Bin}(n,p)$

$$P(N=x) \quad = \quad \begin{pmatrix} n \\ x \end{pmatrix} p^x(1-p)^{n-x} \text{ pour } x = 1,\ldots,n$$

Note $\begin{pmatrix} n \\ x \end{pmatrix} = \frac{n!}{x!(n-x)!}$

espérance	$\mathbb{E}[X] = np$
variance	$\mathrm{Var}(X) = np(1-p)$

- Distribution multinomiale

$$\begin{pmatrix} n \\ k_1\, k_2\, \ldots\, k_m \end{pmatrix} \triangleq \frac{n!}{k_1!k_2!\cdots k_m!}$$

$$P(k_1,k_2,\ldots,k_m) = \begin{pmatrix} k_1\, k_2\, \ldots\, k_m \end{pmatrix} p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m}$$

Combinatoire

- loi multinomiale (cf. ci-dessus) — tirage avec remise
- loi hypergéométrique — tirage sans remise

$$P(A_k) = \frac{\begin{pmatrix} r \\ k \end{pmatrix}\begin{pmatrix} N-r \\ n-k \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}} \qquad k = 0,1,2,\ldots,\min(r,n)$$

Disrtributions courantes

Distribution uniforme: $X \sim \mathcal{U}(a,b)$ espérance $\mu = \mathbb{E}[X] = \frac{a+b}{2}$
variance $\mathrm{Var}(X) = \frac{(b-a)^2}{12}$

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{sinon} \end{cases} \qquad F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & x > b \end{cases}$$

Distribution exponentielle: $X \sim \mathcal{E}(\lambda)$ espérance $\mu = \mathbb{E}[X] = \frac{1}{\lambda}$
variance $\mathrm{Var}(X) = \frac{1}{\lambda^2}$

$$f(x) = \lambda\, e^{-\lambda x} 1_{\{x \geq 0\}} \qquad F(x) = (1 - e^{-\lambda x})\, 1_{\{x \geq 0\}}$$

Distribution Gamma: $X \sim \Gamma(r, \lambda)$ espérance $\mu = \mathbb{E}[X] = \frac{r}{\lambda}$
variance $\mathrm{Var}(X) = \frac{r}{\lambda^2}$

$$f(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} 1_{\{x \geq 0\}}$$

$F(x)$ pas d'expression mathématique simple

Distribution du chi carré: $X \sim \chi^2_{(\nu)}$ équivalent à $\Gamma(\nu/2, 1/2)$
espérance $\mu = \mathbb{E}[X] = \nu$
variance $\mathrm{Var}(X) = 2\nu$

$$f(x) = \frac{2^{-\nu/2}}{\Gamma(\nu/2)} x^{\nu/2-1} e^{-\frac{1}{2}x} 1_{\{x \geq 0\}}$$

Distribution de Student: $T \sim t_{(\nu)}$
espérance $\mu = \mathbb{E}[T] = 0$ si $\nu > 1$, n'existe pas si $\nu = 0$
variance $\mathrm{Var}(T) = \frac{\nu}{\nu+2}$ si $\nu > 2$ sinon $+\infty$

$$f(x) = c(1 + \frac{t^2}{\nu})^{-\frac{n+1}{2}} \text{ avec } c = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n}\Gamma(1/2)\Gamma(\nu/2)}$$

Distribution gaussienne, (distribution normale)

à une dimension: $X \sim \mathcal{N}(\mu, \sigma^2)$

espérance	$\mu = \mathbb{E}[X] = \mu$
variance	$\text{var}(X) = \sigma^2$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)$$

$F(x)$ pas de forme fermée simple

Normale réduite: $Z \sim \mathcal{N}(0,1)$ avec $Z = \frac{X-\mu}{\sigma}$

multivariable: $X \sim \mathcal{N}(\mu, \Sigma)$ $X = (X_1, \ldots, X_d)$

$$p_X(x) - (2\pi)^{-d/2}|\Sigma|^{-1/2} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$$

$$\mathbb{E}X = \mu \qquad \Sigma_{ij} = \mathbb{E}[(X_i - \mathbb{E}X_i)(X_j - \mathbb{E}X_j)]$$

distance de Mahalanobis : r

$$r^2 = (x-\mu)^T\Sigma^{-1}(x-\mu)$$

Distribution géométrique

Variable aléatoire géométrique : Une variable aléatoire géométrique correspond au nombre d'essai de Bernouilli nécessaires avant une première réussite.

Distribution géométrique :
$X \sim \text{Geo}(p)$ avec p la probabilité de réussite

$$P(X = x) = p(1-p)^{x-1} \text{ pour } x = 1, 2, 3, \ldots$$

espérance	$\mu = \mathbb{E}[X] = \frac{1}{p}$
variance	$\text{Var}(X) = \frac{1-p}{p^2}$

Distribution de Poisson

$X \sim \text{Pois}(\lambda)$ avec λ le taux des évènements (taux d'occurences)

$$P(X = x) = f(x) = \frac{\lambda^x \, e^{-\lambda}}{x!} \text{ pour } x = 0, 1, 2, 3, \ldots$$

espérance	$\mathbb{E}[X] = \lambda$
variance	$\text{Var}(X) = \lambda$

Calcul récursif:

$$\frac{f(x)}{f(x-1)} = \frac{\lambda}{x}$$

Laplace—de Moivre

Soit la distribution binomiale $\mathcal{B}in(n, p)$ donnée par

$$P_n(k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad k = 0, 1, \ldots, n$$

On a l'approximation lorsque n est grand par

$$\boxed{P_n(k) \sim \frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{1}{2}\left(\frac{k-np}{\sqrt{np(1-p)}}\right)^2}}$$

Binomiale et Poisson

Lorsque $\lambda = np$ et $n >> n\,p$, la binomiale s'approxime par une distribution de Poisson

$$\binom{n}{x} p^x \, (1-p)^{n-x} \sim \frac{\lambda^x \, e^{-\lambda}}{x!}$$