

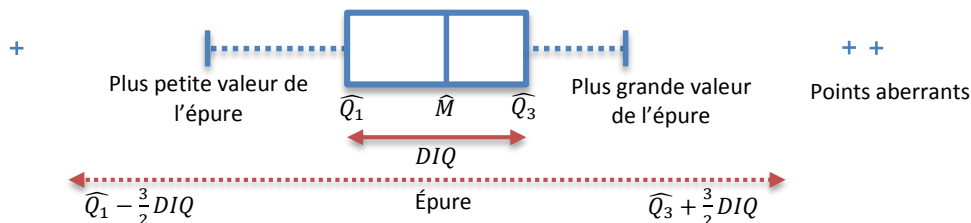
Résumé de Statistiques

Stats – Résumé

I. Statistiques descriptive

- **Quantitative** : Modalités comparables entre elles
 - **Discrète / continue** : modalités dénombrable / indénombrables
- **Qualitative**
 - **Binaire / Multimodale** : deux modalités / plus de 2.
 - **Ordinale (ou non)** : existence d'un ordre

Définition	Formule				Qu	Qu	Qu	Qu	Q		
• Fonction de répartition empirique $\widehat{F}_X$	$\widehat{F}_X(x) = \widehat{F}_l^c + (\widehat{F}_l^c - \widehat{F}_{l-1}^c) \frac{x - x_{l-1}}{x_l - x_{l-1}}$ <div>$\widehat{F}_X(x_l) = \widehat{F}_l^c = \widehat{F}_l - \frac{1}{2}(\widehat{F}_l - \widehat{F}_{l-1})$</div>								X		
• Moyenne $\bar{x}$	Moyenne $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$	Variance empirique biaisée $\hat{S}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2$		Variance débiaisée $S^{*2} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$		x	x	X	X		
• Fractile et quartile	$\widehat{F}_X(\widehat{\phi}_p) = p$	$\widehat{Q}_1 = \widehat{\phi}_{\frac{1}{4}}$	$\widehat{Q}_3 = \widehat{\phi}_{\frac{3}{4}}$	$DIQ = \widehat{Q}_3 - \widehat{Q}_1$	$Etendue = x_{\max} - x_{\min} $			X	X		
• Moment et moment centré	$\widehat{m}_k = \frac{1}{n} \sum_{i=1}^n x_i^k$	$\widehat{\mu}_k = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k$		$\bar{x} = \widehat{m}_1$ Moyenne	$\hat{S}^2 = \widehat{\mu}_2$ Variance	$\frac{\widehat{\mu}_3}{\hat{S}^3}$ Dissymétrie	$\frac{\widehat{\mu}_4}{\hat{S}^4}$ Aplatissement	x	x	X	X
• Dépendance : (X, Y quant.)	$s_{XY}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n} \sum_{i=1}^n (x_i y_i - \bar{x} \bar{y})$ Covariance		$s_X^2 = s_{XX} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ $s_Y^2 = s_{YY} = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$ Variance		$r = \frac{s_{XY}}{s_X s_Y}$ Corrélation $\in [-1; 1]$ $r \approx 0 \Rightarrow x \text{ indep } y$ $r \approx 1 \Rightarrow x \nearrow y \nearrow$ $r \approx -1 \Rightarrow x \nearrow y \searrow$	$a = \frac{cov(X, Y)}{s_X^2}$ $b = \bar{y} - a \bar{x}$					
• Dépendance : (X qual, Y quant.)	$S_Y^2 = \underbrace{\frac{1}{n} \sum_{j \in \Omega_X} n_j (\bar{y}_j - \bar{y})^2}_{S_E^2: \text{variance expliquée}} + \underbrace{\frac{1}{n} \sum_{j \in \Omega_X} n_j s_j^2}_{S_R^2: \text{variance résiduelle}}$				$S_{Y/X}^2 = \frac{\sum_{j \in \Omega_X} n_j (\bar{y}_j - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$ Coef. de détermination						
• Dépendance : (X, Y qual.)	<div>$D_{\chi^2} = \sum_{i \in \Omega_X} \sum_{j \in \Omega_Y} \frac{\left(n_{ij} - \frac{n_{i\bullet} n_{\bullet j}}{n}\right)^2}{\frac{n_{i\bullet} n_{\bullet j}}{n}} = n \sum_{i \in \Omega_X} \sum_{j \in \Omega_Y} \frac{\left(\widehat{\mathbb{P}}_{ij} - \widehat{\mathbb{P}}_i \widehat{\mathbb{P}}_j\right)^2}{\widehat{\mathbb{P}}_i \widehat{\mathbb{P}}_j}$</div> <div>$\Phi^2 = \frac{\chi^2}{n} \quad T = \sqrt{\frac{\Phi^2}{\sqrt{(r-1)(c-1)}}} \quad C = \sqrt{\frac{\Phi^2}{\min(r, c) - 1}}$<div>Coef de Tschuorow<div>Coef de Cramer</div></div></div>										



Résumé de Statistiques

Stats – Résumé

II. Probabilités : variables aléatoires

1. Formules

	Discrète	Continue
Def.	Loi de probabilité $P(\Omega) = 1$ $P_X(x) = P(X = x)$	Fonction de densité $\int_{-\infty}^{+\infty} f(t) dt = 1$
FDR	$F_X(x) = P(X \leq x)$	$F_X(x) = \int_{-\infty}^x f(t) dt \quad F_X = \int f$
Somme $Z = X + Y$	$P(Z = z) = \sum_i P(X = x_i) P_{(X=x_i)}(Y = z - x_i)$	$h(x) = f * g = \int_{-\infty}^{+\infty} f(t)g(x - t) dt$ <p>avec * produit de convolution</p>
Espérance	$E(X) = \sum_{i=1}^n x_i P(X = x_i)$ $E(\varphi(X)) = \sum_{i=1}^n \varphi(x_i) P(X = x_i)$ $E(aX + b) = aE(X) + b$	$E(X) = \int_{-\infty}^{+\infty} t f(t) dt$ $E(\varphi(X)) = \int_{-\infty}^{+\infty} \varphi(t) f(t) dt$
Variance	$V(X) = E\left((X - E(X))^2\right)$ $V(X) = E(X^2) - E(X)^2$ $\sigma(X) = \sqrt{V(X)}$	Propriétés : <ul style="list-style-type: none"> $V(aX + b) = a^2 V(X)$ $\sigma(aX + b) = a \sigma(X)$

2. Lois usuelles

Nom	$P(X = k) / f(x)$	$E(X)$	$V(X)$	Propriétés
Uniforme $U(n)$	$\frac{1}{n}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$	$\Omega = \llbracket 1; n \rrbracket$
Binomiale $\mathcal{B}(n, p)$	$\binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$	$\Omega = \llbracket 0; n \rrbracket$ $n > 50, \begin{cases} p < 0,1 : \mathcal{B} \rightarrow \mathcal{P} \\ p \geq 0,1 : \mathcal{B} \rightarrow \mathcal{N} \end{cases}$
Poisson $\mathcal{P}(\lambda)$	$e^{-\lambda} \frac{\lambda^k}{k!}$	λ	λ	λ grand : $\mathcal{P} \rightarrow \mathcal{N}$ $X_1 + X_2 \sim \mathcal{P}(\lambda_1 + \lambda_2)$
Uniforme $U_{[0;a]}$	$\frac{1}{a}$ si $x \in [0; a]$	$\frac{a}{2}$	$\frac{a^2}{12}$	
Exponentielle $E(\lambda)$	$\lambda e^{-\lambda x}$ si $x > 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	
Normale $\mathcal{N}(\mu, \sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	μ	σ^2	$Z \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

Résumé de Statistiques

Stats – Résumé

III. Estimateur et fonction de vraisemblance

1. Fonction de vraisemblance

$$\mathcal{L}(x_1, \dots, x_n, \theta) = \prod_{i=1}^n f(x_i; \theta) \quad \mathcal{L}(x_1, \dots, x_n, \theta) = \prod_{i=1}^n P(X = x_i; \theta)$$

2. Estimateurs ponctuels

a. Formules, propriétés, théorèmes

$$\hat{\Theta} = T = f(X_1, \dots, X_n)$$

Form./Prop./Th.	Définition
$\hat{\Theta}$ converge $\Leftrightarrow$	$E(\hat{\Theta}) \xrightarrow{n \rightarrow \infty} \theta$ et $V(\hat{\Theta}) \xrightarrow{n \rightarrow \infty} 0$
$\hat{\Theta}_1$ préférable à $\hat{\Theta}_2 \Leftrightarrow$	$\forall \theta \in \mathbb{R}, R(\hat{\Theta}_1, \theta) \leq R(\hat{\Theta}_2, \theta)$
$\hat{\Theta}$ admissible $\Leftrightarrow$	aucun estimateur préférable
$\hat{\Theta}$ exhaustive $\Leftrightarrow$	contient toute l'info apportée par l'échantillon pour connaître θ $\Leftrightarrow \exists g, h \mid \mathcal{L} = g(t, \theta)h(x_1, \dots, x_n)$
Risque :	$R(\hat{\Theta}, \theta) = E[(\hat{\Theta} - \theta)^2] = V(\hat{\Theta}) + (E(\hat{\Theta}) - \theta)^2$
Information de Fisher :	$I_n(\theta) = E\left[\left(\frac{\partial \ln \mathcal{L}}{\partial \theta}\right)^2\right] = -E\left[\frac{\partial^2 \ln \mathcal{L}}{\partial \theta^2}\right] = nI_1(\theta)$
Thm de Cramer Rao :	<ul style="list-style-type: none"> $\hat{u}$ est. sans biais de $u(\theta)$ Support de X indépendant de θ $\frac{\partial \mathcal{L}}{\partial \theta}$ existe, continu, intégrable $\hat{u} \frac{\partial \mathcal{L}}{\partial \theta}$ intégrable $I_n(\theta)$ finie $\Rightarrow V(\hat{u}) \geq BCR = \frac{u'(\theta)^2}{I_n(\theta)}$
Estimateur efficace $\Leftrightarrow$	$\begin{cases} \text{vérifie les cond. de Cramer-Rao} \\ \text{variance} = \text{variance de la BCR} = \text{variance min} \\ \text{vérifie les cond. de Cramer-Rao} \end{cases}$ $\Leftrightarrow \left\{ \begin{array}{l} \exists A \mid \frac{\partial \ln(\mathcal{L})}{\partial \theta} = A(n, \theta)(\hat{u} - u(\theta)) \end{array} \right.$

b. Création d'un estimateur ponctuel

• Méthode du maximum de variance

$$\max_{\hat{\Theta}} \mathcal{L}$$

$$\hat{\Theta}_{MV} = \theta \mid \frac{\partial \ln \mathcal{L}}{\partial \theta} = 0 \quad \text{et} \quad \frac{\partial^2 \ln \mathcal{L}}{\partial \theta^2} < 0$$

Propriétés :

- $\hat{\Theta}_{MV}$ MV de $\theta \Rightarrow u(\hat{\Theta}_{MV})$ MV de $u(\theta)$
- $\hat{\Theta}_{MV}$ asy. sans biais
- $\hat{\Theta}_{MV}$ asy. efficace et asy. gaussien si CCR vérifiées

• Méthode des moments

On utilise les $\widehat{m}_k$ et les $\widehat{\mu}_k$ et on en déduit des estimateurs $\hat{\Theta}_m$ des paramètres.

Propriétés :

- Convergent : $\hat{\Theta}_m \rightarrow \hat{\Theta}$
- Asymptotiquement gaussien : $\sqrt{n}(\hat{\Theta} - \theta) \rightarrow \mathcal{N}(0, \sigma_m^2)$

Résumé de Statistiques

Stats – Résumé

c. Convergence des estimateurs courants et propriétés

X_i i.i.d. avec $E(X_i) = E(X) = \mu$ $V(X_i) = \sigma^2$

$\hat{\theta}$	$E(\hat{\theta})$	$V(\hat{\theta})$	Loi gd nombres	$X \sim \mathcal{N}(\mu, \sigma^2)$
$\bar{X} = \frac{\sum X_i}{n}$	μ	$\frac{\sigma^2}{n}$	$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \rightarrow \mathcal{N}(0,1)$	$\bar{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$
$S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$	$\frac{n-1}{n} \sigma^2$	$\approx \frac{\mu_4 - \sigma^4}{n}$	$\frac{S^2 - \frac{n-1}{n} \sigma^2}{\sqrt{V(S^2)}} \rightarrow \mathcal{N}(0,1)$	$\frac{nS^2}{\sigma^2} \sim \chi_{n-1}^2$
$S^{*2} = \frac{n}{n-1} S^2$	σ^2			

3. Estimation par intervalle de confiance

a. Principe

$$\mathbb{P}(\theta \in [\hat{\theta}_1, \hat{\theta}_2]) = 1 - \alpha$$

Principe

$1 - \alpha$: niveau de confiance

α : risque

$$[\hat{\theta}_1, +\infty[$$

Intervalle de confiance unilatéral

$$[\hat{\theta}_1, \hat{\theta}_2]$$

Intervalle de confiance bilatéral
risque symétrique ou non

b. Construction

Soit un estimateur $\hat{\theta}$ de loi de probabilité $\mathcal{P}$.

1. Construire la fonction pivotale $f(\hat{\theta}, \theta)$ loi de proba. ne dépendant pas de θ
2. Déterminer $I_{1,2}$ tel que $\mathbb{P}(I_1 \leq f(\hat{\theta}, \theta) \leq I_2) = 1 - \alpha$
3. En déduire $\hat{\theta}_{1,2}$ tel que $\mathbb{P}(\hat{\theta}_1 \leq f(\hat{\theta}, \theta) \leq \hat{\theta}_2) = 1 - \alpha$

c. Fonctions pivotales usuelles

θ	S^2	S^{*2}	$\bar{X}$ σ connu	$\bar{X}$ σ estimé
$\hat{\theta}$	$\sigma^2 = V(X)$		$\mu = E(X)$	
$f(\hat{\theta}, \theta)$	$\frac{nS^2}{\sigma^2} \sim \chi_{n-1}^2$	$\frac{(n-1)S^{*2}}{\sigma^2} \sim \chi_{n-1}^2$	$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0,1)$	$\frac{\bar{X} - \mu}{\sqrt{\frac{S^{*2}}{n}}} \sim \mathcal{T}_{n-1}$
Prop.	<ul style="list-style-type: none"> $E(\chi_p^2) = p$ $V(\chi_p^2) = 2p$ $\sqrt{2\chi_p^2} \underset{p>30}{\sim} \mathcal{N}(\sqrt{2p-1}, 1)$ 		<ul style="list-style-type: none"> $E(\mathcal{T}_p) = 0$ $V(\mathcal{T}_p) = \frac{p}{p-2}$ $\mathcal{T}_p \underset{p>30}{\sim} \mathcal{N}\left(0, \frac{p}{p-2}\right)$ 	

IV. Tests statistiques

1. Les hypothèses

$\mathcal{H}_0$: hypothèse nulle (privilégiée, plus vraisemblable / habituelle / prudente / facile à formuler)
 $\mathcal{H}_1$: hypothèse alternative

- Hypothèse simple : $\theta = \theta_0$ (autres param. connus)
- Hypothèse composite : $\theta \in \Omega$ (ou $\theta = \theta_0$ avec param. inconnu(s))

2. Risque

- Risque de 1^{ère} espèce : $\alpha = \mathbb{P}(\mathcal{H}_1 | \mathcal{H}_0 \text{ vraie})$
- Risque de 2^{nde} espèce : $\beta = \mathbb{P}(\mathcal{H}_0 | \mathcal{H}_1 \text{ vraie})$
- Puissance du test : $1 - \beta = \mathbb{P}(\mathcal{H}_1 | \mathcal{H}_1 \text{ vraie})$

Vérité Décision	$\mathcal{H}_0$	$\mathcal{H}_1$
$\mathcal{H}_0$	$1 - \alpha$	β
$\mathcal{H}_1$	α	$1 - \beta$

α fixé, β à minimiser

3. Règle de décision

R_{UPP} est uniformément plus puissant si :

$$\forall \theta, \forall R, (1 - \beta_{R_{UPP}}) \geq (1 - \beta_R)$$

Choix de la statistique T variable de décision : tel que T soit exhaustif et efficace, de loi différente sous $\mathcal{H}_{0/1}$ et de loi connue sous $\mathcal{H}_0$.

Région critique W : ensemble des valeurs de T tel que l'on choisisse $\mathcal{H}_1$. ($\bar{W}$ r. d'acceptation)

$$\mathbb{P}(W | \mathcal{H}_0) = \alpha \quad \mathbb{P}(W | \mathcal{H}_1) = 1 - \beta$$

Résumé de Statistiques

Stats – Résumé

4. Tests paramétriques

	HS-HS	HS-HC		HC-HC	
$\mathcal{H}_0$	$\theta = \theta_0$	$\theta = \theta_0$	$\theta = \theta_0$	$\theta \leq \theta_0$	$\theta = \theta_0$ params inc.
$\mathcal{H}_1$	$\theta = \theta_1 \geq \theta_0$	$\theta \geq \theta_0$ $\Leftrightarrow \theta = \theta_i \geq \theta_0$	$\theta \neq \theta_0$	$\theta > \theta_0$	$\theta \geq \theta_0$ params inc.
W	$W = \{\hat{\theta} \geq A\}$ $W = \left\{ \frac{\mathcal{L}(x_1, \dots, x_n, \theta_1)}{\mathcal{L}(x_1, \dots, x_n, \theta_0)} > k_\alpha \right\}$ RC optimale d'après Neyman-Pearson	$W = \{\hat{\theta} < A_1\} \cup \{\hat{\theta} > A_2\}$		$W = \{\hat{\theta} > A\}$ Si $\frac{\mathcal{L}(x_1, \dots, x_n, \theta')}{\mathcal{L}(x_1, \dots, x_n, \theta)} \nearrow \hat{\theta}$ pour $\theta' > \theta$	basé sur fct. pivotale Même RC que les autres
A	$\mathbb{P}(W \mathcal{H}_0) = \alpha_{\text{fixé}} \rightarrow A$	$\mathbb{P}(W \mathcal{H}_0) = \alpha_{\text{fixé}}$ $\Leftrightarrow \underbrace{\mathbb{P}(\hat{\theta} < A_1 \mathcal{H}_0)}_{=\frac{\alpha}{2}} + \underbrace{\mathbb{P}(\hat{\theta} > A_2 \mathcal{H}_0)}_{=\frac{\alpha}{2}} = \alpha$ $\rightarrow A_1, A_2$		$\mathbb{P}(\hat{\theta} > A \mathcal{H}_0, \text{critique})_{\theta = \theta_0} = \alpha_{\text{fixé}} \rightarrow A$	
UPP	UPP	UPP	Non-UPP	UPP	Non-UPP
α	$\alpha = \alpha_{\text{fixé}}$	$\alpha = \alpha_{\text{fixé}}$	$\alpha \leq \alpha_{\text{fixé}}$	$\alpha \leq \alpha_{\text{fixé}}$	$\alpha \leq \alpha_{\text{fixé}}$
β	$\beta = cst$	$\beta = f(\theta)$	$\beta = f(\theta)$	$\beta = f(\theta)$	$\beta = f(\theta)$

Voir si $\hat{\theta} \in W$ ou pas

5. Tests non-paramétrique

	Comparaison d'échantillons appariés	Tests d'adéquation	
	Test du signe	Test du χ^2	Test de Kolmogorov-Smirnov
$\begin{cases} \mathcal{H}_0 \\ \mathcal{H}_1 \end{cases}$	$\begin{cases} \mathcal{H}_0: p = p_0 = \frac{1}{2} \\ \mathcal{H}_1: p = p_1 \neq p_0 \end{cases}$	$\begin{cases} \mathcal{H}_0: F = F_0 \left(\mathcal{H}_0: \hat{N}_k = N_k = np_k \right) \\ \mathcal{H}_1: F \neq F_0 \left(\mathcal{H}_1: \hat{N}_k \neq N_k \right) \end{cases}$	$\begin{cases} \mathcal{H}_0: F = F_0 \\ \mathcal{H}_1: F \neq F_0 \end{cases}$
T	$Z = \text{count } Y_i > X_i$	$D^2 = \sum_{k=1}^K \frac{(\hat{N}_k - N_k)^2}{N_k}$	$D_n = \max_x F_n^*(x) - F_0(x) $
$f(T) \stackrel{\mathcal{H}_0}{\sim}$	$Z \stackrel{\mathcal{H}_0}{\sim} \mathcal{B}(n, p = p_0)$	$D^2 \stackrel{\mathcal{H}_0}{\sim} \chi_{K-1}^2$	$D_n \stackrel{\mathcal{H}_0}{\sim} D_n$
W	$W = \{Z < A\} \cup \{Z > B\}$ $A = b_{n,p,\frac{\alpha}{2}}$ $B = b_{n,p,1-\frac{\alpha}{2}}$	$W = \{D^2 > A\}$ $A = \chi_{K-1-r,1-\alpha}$	$W = \{D_n > A\}$ $A = d_{n,1-\alpha}$
Rqs	$X_i = Y_i$: on supprime le couple $Y > X \Leftrightarrow p_1 > p_0$	But : X suit une loi déterminée ? K classes / n obs. / r params. inconnus sous $\mathcal{H}_0$ $N_k / \hat{N}_k$ nombre de valeurs (théorique / mesurées) dans la classe k Si $N_k < 5$, regrouper des classes	

Résumé de Statistiques

Stats – Résumé

	Test séquentiel	Comparaison d'échantillons non appariés		
		Gaussiens		Distribution inconnue
		Test de Fisher (égalité des variances)	Test de Student (égalité des moyennes à variance égale)	Test non-paramétrique de Wilcoxon
$\begin{cases} \mathcal{H}_0 \\ \mathcal{H}_1 \end{cases}$	$\begin{cases} \mathcal{H}_0: p = p_0 \\ \mathcal{H}_1: p = p_1 > p_0 \end{cases}$	$\begin{cases} \mathcal{H}_0: \sigma_X^2 = \sigma_Y^2 \\ \mathcal{H}_1: \sigma_X^2 \neq \sigma_Y^2 \end{cases}$	$\begin{cases} \mathcal{H}_0: \mu_X = \mu_Y \\ \mathcal{H}_1: \mu_X \neq \mu_Y \end{cases}$	$\begin{cases} \mathcal{H}_0: F_X = F_Y \\ \mathcal{H}_1: F_X > F_Y \end{cases}$
T	$Z_i = \ln \frac{\mathcal{L}(X_1, \dots, X_n, p_1)}{\mathcal{L}(X_i, p_0)}$	$\frac{(n-1)}{\sigma_X^2} S_X^{*2} \sim \chi_{n-1}^2 \quad \frac{(m-1)}{\sigma_Y^2} S_Y^{*2} \sim \chi_{m-1}^2$	$\bar{X}, \bar{Y}$	$rangs = sort([X \ Y])$ $W_X = \sum \text{rang } X_i$
$f(T) \stackrel{\mathcal{H}_0}{\sim}$		$\frac{S_X^{*2}}{S_Y^{*2}} \stackrel{\mathcal{H}_0}{\sim} \mathcal{F}_{n-1, m-1}$	$\frac{\bar{X} - \bar{Y}}{S^* \sqrt{\frac{1}{n} + \frac{1}{m}}} \stackrel{\mathcal{H}_0}{\sim} \mathcal{T}_{n+m-2}$	$\frac{W_X - \mu_W}{\sigma_W} \stackrel{\mathcal{H}_0}{\rightarrow} \mathcal{N}(0,1)$ $\mu_W = \frac{n(n+m+1)}{2}$ $\sigma_W^2 = \frac{nm(n+m+1)}{12}$
W	$\begin{cases} \sum_{i=1}^n Z_i \geq \ln \frac{1}{\alpha} & \mathcal{H}_1 \\ \sum_{i=1}^n Z_i \leq \ln \beta & \mathcal{H}_0 \\ \text{sinon} & n = n+1 \end{cases}$	$W = \{T < A\} \cup \{T > B\}$ $A = f_{n-1, m-1, \frac{\alpha}{2}}$ $B = f_{n-1, m-1, 1-\frac{\alpha}{2}}$	$W = \{T < A\} \cup \{T > B\}$ $A = t_{n+m-2, \frac{\alpha}{2}}$ $B = t_{n+m-2, 1-\frac{\alpha}{2}}$	$W = \{T < A\}$ $A = n_{1-\alpha}$
Rqs		$S^{*2} = \frac{(n-1)S_X^{*2} + (m-1)S_Y^{*2}}{n+m-2}$ $V_1 \sim \chi_n^2 \quad V_2 \sim \chi_m^2 \quad \frac{\frac{V_1}{n}}{\frac{V_2}{m}} \sim \mathcal{F}_{n,m}$ $f_{n,m,\alpha} = f_{m,n,1-\alpha}$		$F_X > F_Y \Leftrightarrow X < Y$ Si $X_i = X_j$, rang = rang moyen

Résumé de Statistiques

Stats – Résumé

Autres

6. Formules

$$\int_0^{+\infty} x^n e^{-\alpha x} dx = \frac{n!}{\alpha^{n+1}}$$

7. Définition de probabilité

Ω ensemble fini P proba sur Ω $P : \Omega \rightarrow [0, 1]$ vérifie :

- $P(\Omega) = 1$
- $\forall A, B \in \mathcal{P}(\Omega)^2 \quad A \cap B = \emptyset \quad P(A \cup B) = P(A) + P(B)$

Propriétés

- $P(\emptyset) = 0$
- $P(A^c) = 1 - P(A)$
- $P(A - B) = P(A) - P(A \cap B)$
- Si $A \subset B$ alors $P(A) \leq P(B)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

8. Formules de probabilité

$$\binom{n}{p} = \frac{n!}{p!(n-p)!}$$

Formule de Poincaré ou formule d'inclusion-exclusion :

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k-1} \left(\sum_{\substack{i_1 \dots i_k \\ 1 \leq i_1 < \dots < i_k \leq n}} P\left(\bigcap_{j=i_1}^{i_k} A_j\right) \right)$$

Probas conditionnelles :

$$P_A(B) = \frac{P(A \cap B)}{P(A)}$$

Probas composées :

$$P\left(\bigcap_{i=1}^n A_i\right) = \prod_{i=1}^n P_{\bigcap_{j=1}^{i-1} A_j}(A_i)$$

Probas totales :

$$P(B) = \sum_{i=1}^n P(A_i) P_{A_i}(B)$$

Tchebychev :

$$P(|X - E(X)| > k\sigma) \leq \frac{1}{k^2}$$

$$V(X + Y) = V(X) + V(Y) + 2cov(X, Y)$$

$$cov(X, Y) = E(XY) - E(X)E(Y) = E((X - E(X))(Y - E(Y)))$$

9. Statistique exhaustive : Théorème de Darmois

X VA dont le domaine ne dépasse pas de θ $(X_1, \dots, X_n)$ échantillon i.i.d.

$$f(x) = e^{a(x)\alpha(\theta)+b(x)+\beta(\theta)} \Rightarrow \text{stat. suffisante}$$

$$a \text{ bijective et cont. différentiable} \Rightarrow T = \sum a(X_i) \text{ stat. exhaustive}$$

10. Théorème de Huygens

$$\forall a \in \mathbb{R}, \sum_{i=1}^n (X_i - a)^2 = \sum_{i=1}^n (X_i - \bar{X})^2 + n(\bar{X} - a)^2$$