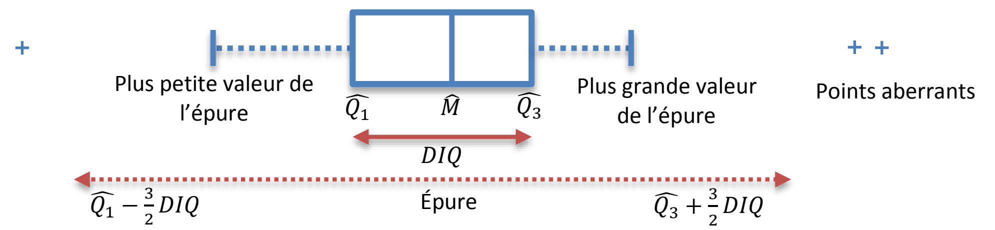
Résumé de Statistiques

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	Qualit. binaire	Qualit. multimod	Qualit. ordinale	Quantit. discrète	Quantit. continue
Fonction de répartition empirique $\widehat{F_X}$	$\widehat{F_X}(x) = \widehat{F_i^c} + (\widehat{F_i^c} - \widehat{F_{i-1}^c})\frac{x - x_{i-1}}{x_i - x_{i-1}}$ $\boxed{\widehat{F_X}(x_i) = \widehat{F_i^c} = \widehat{F_i} - \frac{1}{2}(\widehat{F_i} - \widehat{F_{i-1}})}$					X
Moyenne $\bar{x}$	Moyenne: $\boxed{\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i}$ Variance empirique biaisée: $\boxed{\hat{S}^2 = \frac{1}{n}\sum_{i=n}^{n}(x_i - \bar{x})^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2}$ Variance débiaisée: $\boxed{S^{*2} = \frac{1}{n}\sum_{i=n}^{n}(x_i - \bar{x})^2}$		x	x	X	X
Fractile et quartile	$\widehat{F_X}(\hat{\phi}_p) = p$; $\widehat{Q_1} = \hat{\phi}_{\frac{1}{4}}$; $\widehat{Q_3} = \hat{\phi}_{\frac{3}{4}}$; $DIQ = \widehat{Q_3} - \widehat{Q_1}$; $Etendue = \lvert x_{\max} - x_{\min}\rvert$				X	X
Moment et moment centré	$\widehat{m_k} = \frac{1}{n}\sum_{i=1}^{n} x_i^k$; $\widehat{\mu_k} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^k$ $\underset{\text{Moyenne}}{\bar{x} = \widehat{m_1}}$ $\underset{\text{Variance}}{\hat{S}^2 = \widehat{\mu_2}}$ $\underset{\text{Dissymétrie}}{\frac{\widehat{\mu_3}}{\hat{S}^3}}$ $\underset{\text{Aplatissement}}{\frac{\widehat{\mu_4}}{\hat{S}^4}}$		x	x	X	X
Dépendance : (X, Y quant.)	$\boxed{s_{XY}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n}\sum_{i=1}^{n}(x_i y_i - \bar{x}\bar{y})}$ Covariance $\boxed{s_X^2 = s_{XX} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \quad s_Y^2 = s_{YY} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2}$ Variance $\boxed{r = \frac{s_{XY}}{s_X s_Y}}$ Corrélation $\in [-1;1]$ $r \approx 0 \Rightarrow x$ 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.)	$S_Y^2 = \underbrace{\frac{1}{n}\sum_{j \in \Omega_X} n_j(\bar{y}_j - \bar{y})^2}_{s_E^2:\text{variance expliquée}} + \underbrace{\frac{1}{n}\sum_{j \in \Omega_X} n_j s_j^2}_{s_R^2:\text{variance résiduelle}}$ $\boxed{S_{Y/X}^2 = \frac{\sum_{j \in \Omega_X} n_j(\bar{y}_j - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$ Coef. de détermination					
Dépendance : (X, Y qual.)	$\boxed{D_{\chi^2} = \sum_{i \in \Omega_X}\sum_{j \in \Omega_Y}\frac{\left(n_{ij} - \frac{n_{i\bullet}n_{\bullet j}}{n}\right)^2}{\frac{n_{i\bullet}n_{\bullet j}}{n}} = n\sum_{i \in \Omega_X}\sum_{j \in \Omega_Y}\frac{\left(\overbrace{\widehat{\mathbb{P}}_{ij}}^{\text{eff obs}} - \overbrace{\widehat{\mathbb{P}}_i\widehat{\mathbb{P}}_j}^{\text{eff th}}\right)^2}{\widehat{\mathbb{P}}_i\widehat{\mathbb{P}}_j}}$ $\Phi^2 = \frac{\chi^2}{n}$; $T = \sqrt{\frac{\Phi^2}{\sqrt{(r-1)(c-1)}}}$ Coef de Tschuorow ; $C = \sqrt{\frac{\Phi^2}{\min(r,c) - 1}}$ Coef de Cramer					



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 $Z = X + Y$	$\boxed{P(Z = z) = \sum_i P(X = x_i)\, P_{(X=x_i)}(Y = z - x_i)}$	$\boxed{h(x) = f * g = \int_{-\infty}^{+\infty} f(t)g(x-t)\,dt}$ avec $*$ produit de convolution
Espérance	$\boxed{E(X) = \sum_{i=1}^{n} x_i\, P(X = x_i)}$ $E(\varphi(X)) = \sum_{i=1}^{n} \varphi(x_i)\, P(X = x_i)$ $E(aX + b) = aE(X) + b$	$\boxed{E(X) = \int_{-\infty}^{+\infty} t\, f(t)\,dt}$ $\boxed{E(\varphi(X)) = \int_{-\infty}^{+\infty} \varphi(t)\, f(t)\,dt}$
Variance	$\boxed{V(X) = E\left(\left(X - E(X)\right)^2\right)}$ $\boxed{V(X) = E(X^2) - E(X)^2}$ $\boxed{\sigma(X) = \sqrt{V(X)}}$	**Propriétés :** • $V(aX + b) = a^2 V(X)$ • $\sigma(aX + b) = \lvert a\rvert\sigma(X)$

2. Lois usuelles

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

III. Estimateur et fonction de vraisemblance

1. Fonction de vraisemblance

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

2. Estimateurs ponctuels

a. Formules, propriétés, théorèmes

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

Form./Prop./Th.	Définition
$\widehat{\Theta}$ converge ⇔	$E(\widehat{\Theta}) \xrightarrow{n\to\infty} \theta$ et $V(\widehat{\Theta}) \xrightarrow{n\to\infty} 0$
$\widehat{\Theta}_1$ préférable à $\widehat{\Theta}_2$ ⇔	$\forall\, \theta \in \mathbb{R}, R(\widehat{\Theta}_1, \theta) \leq R(\widehat{\Theta}_2, \theta)$
$\widehat{\Theta}$ admissible ⇔	aucun estimateur préférable
$\widehat{\Theta}$ exhaustive ⇔	contient toute l'info apportée par l'échantillon pour connaitre θ $\Leftrightarrow \exists\, g, h \mid \mathcal{L} = g(t,\theta)h(x_1, \dots, x_n)$
Risque :	$R(\widehat{\Theta}, \theta) = E\left[\left(\widehat{\Theta} - \theta\right)^2\right] = V(\widehat{\Theta}) + \left(E(\widehat{\Theta}) - \theta\right)^2$
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 :	• $\hat{u}$ 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 $\Rightarrow \quad V(\hat{u}) \geq BCR = \frac{u'(\theta)^2}{I_n(\theta)}$
Estimateur efficace ⇔	$\begin{cases} \text{vérifie les cond. de Cramer-Rao} \\ \text{variance = variance de la BCR = variance min} \end{cases}$ $\Leftrightarrow \begin{cases} \text{vérifie les cond. de Cramer-Rao} \\ \exists\, A \mid \frac{\partial\, ln(L)}{\partial \theta} = A(n,\theta)\left(\hat{u} - u(\theta)\right) \end{cases}$

b. Création d'un estimateur ponctuel

- ***Méthode du maximum de variance***

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

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

Propriétés :

- $\widehat{\Theta}_{MV}$ MV de $\theta \Rightarrow u(\widehat{\Theta}_{MV})$ MV de $u(\theta)$
- $\widehat{\Theta}_{MV}$ asy. sans biais
- $\widehat{\Theta}_{MV}$ 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 $\widehat{\Theta}_m$ des paramètres.

Propriétés :

- Convergent : $\widehat{\Theta}_m \to \widehat{\Theta}$
- Asymptotiquement gaussien : $\sqrt{n}\left(\widehat{\Theta} - \theta\right) \to \mathcal{N}(0, \sigma_m^2)$

c. Convergence des estimateurs courants et propriétés

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

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

3. Estimation par intervalle de confiance

a. Principe

$$\mathbb{P}\left(\theta \in [\widehat{\Theta}_1, \widehat{\Theta}_2]\right) = 1 - \alpha$$

Principe

$1 - \alpha$: niveau de confiance

α : risque

$$[\widehat{\Theta}_1, +\infty[$$

Intervalle de confiance unilatéral

$$[\widehat{\Theta}_1, \widehat{\Theta}_2]$$

Intervalle de confiance bilatéral
risque symétrique ou non

b. Construction

Soit un estimateur $\widehat{\Theta}$ de loi de probabilité $\mathcal{P}$.

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

c. Fonctions pivotales usuelles

$\boldsymbol{\theta}$	$\boldsymbol{S^2}$	$\boldsymbol{S^{*2}}$	$\boldsymbol{\bar{X}}$ **σ connu**	$\boldsymbol{\bar{X}}$ **σ estimé**
$\widehat{\Theta}$	$\sigma^2 = V(X)$		$\mu = E(X)$	
$f(\widehat{\Theta}, \theta)$	$\frac{nS^2}{\sigma^2} \sim \chi^2_{n-1}$	$\frac{(n-1)S^{*2}}{\sigma^2} \sim \chi^2_{n-1}$	$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0,1)$	$\frac{\bar{X} - \mu}{\sqrt{\frac{S^{*2}}{n}}} \sim \mathcal{T}_{n-1}$
Prop.	• $E(\chi^2_p) = p$ • $V(\chi^2_p) = 2p$ • $\sqrt{2\chi^2_p} \underset{p>30}{\sim} \mathcal{N}(\sqrt{2p-1}, 1)$		• $E(\mathcal{T}_p) = 0$ • $V(\mathcal{T}_p) = \frac{p}{p-2}$ • $\mathcal{T}_p \underset{p>30}{\sim} \mathcal{N}\left(0, \frac{p}{p-2}\right)$	

IV. Tests statistiques

1. Les hypothèses

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

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

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

α fixé, β à minimiser

3. Règle de décision

R_{UPP} est uniformément plus puissant si :

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

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 choisi $\mathcal{H}_1$. ($\overline{W}$ r. d'acceptation)

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

4. Tests paramétriques

	HS-HS	HS-HC		HC-HC	
$\boldsymbol{\mathcal{H}_0}$	$\theta = \theta_0$	$\theta = \theta_0$	$\theta = \theta_0$	$\theta \leq \theta_0$	$\theta = \theta_0$ params inc.
$\boldsymbol{\mathcal{H}_1}$	$\theta = \theta_1 \gtrless \theta_0$	$\theta \gtrless \theta_0$ $\Leftrightarrow \theta = \theta_i \gtrless \theta_0$	$\theta \neq \theta_0$	$\theta > \theta_0$	$\theta \gtrless \theta_0$ params inc.
$\boldsymbol{W}$	$W = \{\hat{\theta} \gtrless A\}$ $W = \left\{\frac{\mathcal{L}(x_1, \dots, x_n, \theta_1)}{\mathcal{L}(x_1, \dots, x_n, \theta_0)} > k_\alpha\right\}$ 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
$\boldsymbol{A}$	$\mathbb{P}(\mathrm{W}\|\mathcal{H}_0) = \alpha_{\text{fixé}} \rightarrow A$		$\mathbb{P}(\mathrm{W}\|\mathcal{H}_0) = \alpha_{\text{fixé}}$ $\Leftrightarrow \underbrace{\mathbb{P}(\hat{\theta} < A_1 \| \mathcal{H}_0)}_{=\frac{\alpha}{2}}$ $+ \underbrace{\mathbb{P}(\hat{\theta} > A_2 \| \mathcal{H}_0)}_{=\frac{\alpha}{2}} = \alpha$ $\rightarrow A_1, A_2$	$\mathbb{P}\left(\hat{\theta} > A \middle\| \begin{matrix}\mathcal{H}_{0,\text{ critique}} \\ \theta = \theta_0\end{matrix}\right)$ $= \alpha_{\text{fixé}} \rightarrow A$	
UPP	UPP	UPP	Non-UPP	UPP	Non-UPP
$\boldsymbol{\alpha}$	$\alpha = \alpha_{fixé}$	$\alpha = \alpha_{fixé}$	$\alpha \leq \alpha_{fixé}$	$\alpha \leq \alpha_{fixé}$	$\alpha \leq \alpha_{fixé}$
$\boldsymbol{\beta}$	$\beta = cst$	$\beta = f(\theta)$	$\beta = f(\theta)$	$\beta = f(\theta)$	$\beta = f(\theta)$

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**
$\begin{cases}\boldsymbol{\mathcal{H}_0} \\ \boldsymbol{\mathcal{H}_1}\end{cases}$	$\begin{cases}\mathcal{H}_0 : p = p_0 = \frac{1}{2} \\ \mathcal{H}_0 : p = p_1 \neq p_0\end{cases}$	$\begin{cases}\mathcal{H}_0 : F = F_0 \\ \mathcal{H}_1 : F \neq F_0\end{cases} \left(\begin{matrix}\mathcal{H}_0 : \widehat{N}_k = N_k = np_k \\ \mathcal{H}_1 : \widehat{N}_k \neq N_k\end{matrix}\right)$	$\begin{cases}\mathcal{H}_0 : F = F_0 \\ \mathcal{H}_1 : F \neq F_0\end{cases}$
$\boldsymbol{T}$	$Z = \text{count}\, Y_i > X_i$	$D^2 = \sum_{k=1}^{K} \frac{(\widehat{N}_k - N_k)^2}{N_k}$	$D_n = \max_x \|F_n^*(x) - F_0(x)\|$
$\boldsymbol{f(T)} \overset{\boldsymbol{\mathcal{H}_0}}{\sim}$	$Z \overset{\mathcal{H}_0}{\sim} \mathcal{B}(n, p = p_0)$	$D^2 \overset{\mathcal{H}_0}{\sim} \chi^2_{K-1}$	$D_n \overset{\mathcal{H}_0}{\sim} D_n$
$\boldsymbol{W}$	$W = \{Z < A\} \cup \{Z > B\}$ $A = b_{n,p,\frac{\alpha}{2}}$ $B = b_{n,p,1-\frac{\alpha}{2}}$	$W = \{D^2 > A\}$ $A = \chi_{K-1-r,1-\alpha}$	$W = \{D_n > A\}$ $A = d_{n,1-\alpha}$
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 $\mathcal{H}_0$ $N_k / \widehat{N}_k$ nombre de valeurs (théorique / mesurées) dans la classe k Si $N_k < 5$, regrouper des classes	

	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**
$\begin{cases}\boldsymbol{\mathcal{H}_0}\\ \boldsymbol{\mathcal{H}_1}\end{cases}$	$\begin{cases}\mathcal{H}_0: p = p_0\\ \mathcal{H}_1: p = p_1 > p_0\end{cases}$	$\begin{cases}\mathcal{H}_0: \sigma_X^2 = \sigma_Y^2\\ \mathcal{H}_1: \sigma_X^2 \neq \sigma_Y^2\end{cases}$	$\begin{cases}\mathcal{H}_0: \mu_X = \mu_Y\\ \mathcal{H}_1: \mu_X \neq \mu_Y\end{cases}$	$\begin{cases}\mathcal{H}_0: F_X = F_Y\\ \mathcal{H}_1: F_X > F_Y\end{cases}$
$\boldsymbol{T}$	$Z_i = \ln \frac{\mathcal{L}(X_1, \dots, X_n, p_1)}{\mathcal{L}(X_i, p_0)}$	$\frac{(n-1)}{\sigma_X^2} S_X^{*^2} \sim \chi_{n-1}^2 \quad \frac{(m-1)}{\sigma_Y^2} S_Y^{*^2} \sim \chi_{m-1}^2$	$\bar{X}, \bar{Y}$	$rangs = sort([X\ Y])$ $W_X = \sum \operatorname{rang} X_i$
$\boldsymbol{f(T)} \overset{\boldsymbol{\mathcal{H}_0}}{\sim}$		$\frac{S_X^{*^2}}{S_Y^{*^2}} \overset{\mathcal{H}_0}{\sim} \mathcal{F}_{n-1,m-1}$	$\frac{\bar{X} - \bar{Y}}{S^* \sqrt{\frac{1}{n} + \frac{1}{m}}} \overset{\mathcal{H}_0}{\sim} \mathcal{T}_{n+m-2}$	$\frac{W_X - \mu_W}{\sigma_W} \overset{\mathcal{H}_0}{\rightarrow} \mathcal{N}(0,1)$ $\mu_W = \frac{n(n+m+1)}{2}$ $\sigma_W^2 = \frac{nm(n+m+1)}{12}$
$\boldsymbol{W}$	$\begin{cases}\sum_{i=1}^{n} Z_i \geq \ln \frac{1}{\alpha} & \mathcal{H}_1\\ \sum_{i=1}^{n} Z_i \leq \ln \beta & \mathcal{H}_0\\ \text{sinon} & n = n+1\end{cases}$	$W = \{T < A\} \cup \{T > B\}$ $A = f_{n-1,m-1,\frac{\alpha}{2}}$ $B = f_{n-1,m-1,1-\frac{\alpha}{2}}$	$W = \{T < A\} \cup \{T > B\}$ $A = t_{n+m-2,\frac{\alpha}{2}}$ $B = t_{n+m-2,1-\frac{\alpha}{2}}$	$W = \{T < A\}$ $A = n_{1-\alpha}$
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

Autres

6. Formules

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

7. Définition de probabilité

Ω ensemble fini $\qquad$ P proba sur Ω $\qquad$ $P : \Omega \to [0,1]$ verifie :

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

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 :

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

Probas conditionnelles :

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

Probas composées :

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

Probas totales :

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

Tchebychev :

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

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

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

9. Statistique exhaustive : Théorème de Darmois

X VA dont le domaine ne dep pas de θ $\qquad$ $(X_1, \dots, X_n)$ 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

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