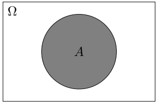
Rappels Chapitre 1 - Introduction aux Probabilités

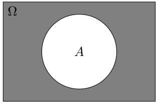
Espace échantillon

$\Omega = S = U$



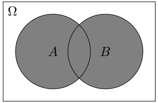
Évènement

A



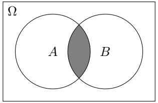
Complémentaire

$\overline{A}$



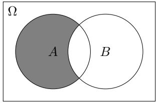
Union (ou)

$A \cup B = \overline{(\overline{A} \cap \overline{B})}$



Intersection (et)

$A \cap B = \overline{(\overline{A} \cup \overline{B})}$



Différence (sans)

$A \setminus B = A \cap \overline{B}$

Cas général

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

A et B disjoints $(A \cap B = \emptyset)$

$\Rightarrow P(A \cup B) = P(A) + P(B)$

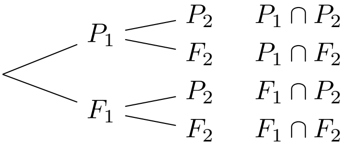
Conditionnelle (sachant)

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

A et B indép.

$\Rightarrow P(A|B) = P(A)$

$\Rightarrow P(A \cap B) = P(A)P(B)$



Rappels Chapitre 1 - Introduction aux Probabilités

- $P(A) = \sum_i^n P(B_i)\, P(A|B_i)$ où $B_1, ..., B_n$: partition de Ω

- Série $\Rightarrow P(F) = P(C_1 \cap ... \cap C_n)$, Indép. $\Rightarrow P(F) = P(C_1) \cdot ... \cdot P(C_n)$

  Paral. $\Rightarrow P(F) = P(C_1 \cup ... \cup C_n)$, Indép. $\Rightarrow P(F) = 1 - (1 - P(C_1)) \cdot ... \cdot (1 - P(C_n))$

- Permutations (d'ordre) de n objets distinguables: $n! = n \cdot (n-1) \cdot (n-2) \cdot ... \cdot 3 \cdot 2 \cdot 1$

  Arrangements (ordonnés) de r objets parmi n distinguables: $\frac{n!}{(n-r)!} = n \cdot ... \cdot (n-r+1)$

  Combinaisons (non-ordonnées) de r objets parmi n distinguables: $\frac{n!}{r!(n-r)!} =: \binom{n}{r}$

Rappels Chapitre 2 - Variables Aléatoires

Discret

- $R_X = \{x \in \mathbb{R} : p_X(x) > 0\} = \{x_1, ..., x_n\}$
- $p_X(x) = P(X = x)$
- $F_X(x) = P(X \leq x) = \sum_{x_i \leq x} p_X(x_i)$

  $P(a < X \leq b) = F_X(b) - F_X(a)$
- $E(g(X)) = \sum_{x \in R_X} g(x) \cdot p_X(x)$

  $\mu_X = E(X) = \sum_{x \in R_X} x \cdot p_X(x)$

  $\sigma_X^2 = E((X - \mu_X)^2) = E(X^2) - (\mu_X)^2$

Continu

- $R_X = \{x \in \mathbb{R} : f_X(x) > 0\} = [a, b]$
- $f_X(x) = \frac{\mathrm{d}}{\mathrm{d}x} P(X \leq x) = \frac{\mathrm{d}}{\mathrm{d}x} F_X(X)$
- $F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(x')\mathrm{d}x'$

  $P(a \leq X \leq b) = F_X(b) - F_X(a)$
- $E(g(X)) = \int_{-\infty}^{\infty} g(x) \cdot f_X(x)\mathrm{d}x$

  $\mu_X = E(X) = \int_{-\infty}^{\infty} x \cdot f_X(x)\mathrm{d}x$

  $\sigma_X^2 = E((X - \mu_X)^2) = E(X^2) - (\mu_X)^2$

- Y discrète avec $Y = H(X)$ et X continue $\Rightarrow p_Y(y) = \int_{\{x : H(x) = y\}} f_X(x)\mathrm{d}x$

Rappels Chapitre 3 - Probabilités Conjointes

- Loi conjointe de X, Y : $p(x,y) = P((X=x)\cap(Y=y))$
- Loi marginale de Y : $p_Y(y) = \sum_{x\in R_X} p(x,y)$
- Loi conditionnelle de Y sachant X : $p_{Y|X=x}(y) = \dfrac{p(x,y)}{p_X(x)}$ $\quad P(A|B) = \dfrac{P(A\cap B)}{P(B)}$
- Covariance : $Cov(X,Y) = E(XY) - E(X)E(Y) = \sum_{x\in R_X}\sum_{y\in R_Y} xy\cdot p(x,y) - \mu_X\mu_y$

  Si X et Y indép. $\Rightarrow Cov(X,Y) = 0 \;\Rightarrow E(XY) = E(X)E(Y)$
- Coefficient de corrélation : $\rho = \dfrac{Cov(X,Y)}{\sigma_X\sigma_Y}$
- Si $Y = a_0 + \sum_{i=1}^{n} a_i X_i$, alors $E(Y) = a_0 + \sum_{i=1}^{n} a_i E(X_i)$

  et $Var(Y) = \sum_{i=1}^{n} a_i^2 Var(X_i) + \sum\sum_{i\neq j} a_i a_j Cov(X_i, X_j)$

Rappels Chapitre 4 - Lois Discrètes

- Bernouilli : *expérience binaire avec probabilité de succès p*

$$X \sim \text{Bernouilli}(p), \quad E(X) = p, \quad Var(X) = p(1-p)$$

$$R_X = \{1 \text{ (succès)}, 0 \text{ (échec)}\}, \quad p_X(1) = p, \quad p_X(0) = 1-p$$

- Binomiale : *nombre de succès après n répétitions d'une expérience binaire avec probabilité de succès p*

$$X \sim \text{B}(n,p), \quad E(X) = np, \quad Var(X) = np(1-p)$$

$$R_X = \{0, 1, 2, ..., n\}, \quad p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$p_X(x) \approx p_{\text{Poi(np)}}(x) \text{ si } n \geq 100,\ p < 0.1,\ np \leq 10$$

- Géométrique : *nombre de répétitions pour avoir un premier succès (de proba. p)*

$$X \sim \text{G}(p), \quad E(X) = \frac{1}{p}, \quad Var(X) = \frac{1-p}{p^2}$$

$$R_X = \{1, 2, 3, ...\}, \quad p_X(x) = p(1-p)^{x-1}, \quad F_X(x) = 1 - (1-p)^x$$

absence de mémoire : $P(X > t + s | X > t) = P(X > s)$

Rappels Chapitre 4 - Lois Discrètes

- Hypergéométrique : *on tire n objets d'un ensemble de N objets dont D sont particuliers, X est le nombre d'objets particuliers tirés*

$$X \sim \mathrm{H}(n, N, D), \quad E(X) = n\frac{D}{N}, \quad Var(X) = n\frac{D}{N}\frac{N-D}{N}\frac{N-n}{N-1}$$

$$R_X = \{\max\{0, n-(N-D)\}, \ldots, \min\{n, D\}\}, \quad p_X(x) = \frac{\binom{D}{x}\binom{N-D}{n-x}}{\binom{N}{n}}$$

$$p_X(x) \approx p_{\mathrm{B}\left(n, \frac{D}{N}\right)}(x) \text{ si } \frac{n}{N} \leq \frac{1}{10}$$

- Poisson : *nombre d'évènements réalisés sur une mesure m (de temps, de distance, de surface, etc.) avec un taux λ par unité de mesure*

$$X \sim \mathrm{Poi}(c) \quad \text{où } c = \lambda m, \quad E(X) = c, \quad Var(X) = c$$

$$R_X = \{0, 1, 2, 3, \ldots\}, \quad p_X(x) = \frac{e^{-c}c^x}{x!}$$

Rappels Chapitre 5 - Lois Continues

- Uniforme : *une valeur dans un intervalle* $[\alpha, \beta]$ *où chacune des valeurs possibles a la même probabilité*

  $X \sim \mathrm{U}(\alpha, \beta), \quad E(X) = \frac{\alpha+\beta}{2}, \quad Var(X) = \frac{(\beta-\alpha)^2}{12}$

  $$R_X = [\alpha, \beta], \quad f_X(x) = \frac{1}{\beta-\alpha}, \quad F_X(x) = \begin{cases} 0 & \text{si } x < \alpha \\ \frac{x-\alpha}{\beta-\alpha} & \text{si } x \in [\alpha, \beta] \\ 1 & \text{si } x > \beta \end{cases}$$

- Exponentielle : *temps d'attente avant le premier évènement d'un processus de Poisson de taux* λ *par unité de temps*

  $X \sim \mathrm{Exp}(\lambda), \quad E(X) = \frac{1}{\lambda}, \quad Var(X) = \frac{1}{\lambda^2}$

  $R_X = [0, \infty[, \quad f_X(x) = \lambda e^{-\lambda x}, \quad F_X(x) = 1 - e^{-\lambda x}$

  absence de mémoire : $P(X > t + s | X > t) = P(X > s)$

Rappels Chapitre 6 - Loi Normale

- Normale : *distribution de modélisation naturelle pour les grands nombres*

  $X \sim \mathrm{N}(\mu, \sigma^2), \quad E(X) = \mu, \quad Var(X) = \sigma^2$

  $R_X =]-\infty, \infty[, \quad f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad F_X(x) = \Phi(\frac{x-\mu}{\sigma})$

  où $\Phi(-z) = 1 - \Phi(z)$

  additivité : $N_i \sim \mathrm{N}(\mu_i, \sigma_i^2)$ indép. $\Rightarrow a_0 + \sum a_i N_i \sim \mathrm{N}(a_0 + \sum a_i \mu_i, \sum a_i^2 \sigma_i^2)$

  T.C.L : X_i i.i.d. et n grand $\Rightarrow \sum X_i \underset{approx.}{\sim} \mathrm{N}(n\mu_X, n\sigma_X^2)$

  $\Rightarrow \bar{X} = \frac{\sum X_i}{n} \underset{approx.}{\sim} \mathrm{N}(\mu_X, \frac{\sigma_X^2}{n})$

Rappels Chapitre 7 - Statistiques Descriptives

- Moyenne de l'échantillon $X_1, ..., X_n$: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
- Médiane de l'échantillon $X_1, ..., X_n$: $X_{0.5}$ *tel que la moitié des valeurs lui sont inférieures*
- Mode de l'échantillon $X_1, ..., X_n$: M_0, la valeur la plus fréquente de l'échantillon
- Variance de l'échantillon $X_1, ..., X_n$:
$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n \cdot \bar{X}^2\right)$$
- Écart interquantile de l'échantillon $X_1, ..., X_n$: $IQR = X_{0.75} - X_{0.25}$

Rappels Chapitre 8 - Échantillons Aléatoires et Lois d'Échantillonnage

- Chi-deux : *une loi* χ^2_k *à* k *degrés de liberté a la distribution d'une somme de* k *normales centrées réduites* $Z_i \sim N(0,1)$ *indép. :* $\chi^2_k = \sum_{i=1}^k Z_i^2$

  $X \sim \chi^2_k, \quad E(X) = k, \quad Var(X) = 2k, \quad \text{quantile } P(\chi^2_k > \chi^2_{\alpha;k}) = \alpha$

  additivité : $\chi^2_{k_i}$ indép. $\Rightarrow \sum \chi^2_{k_i} \sim \chi^2_k$ où $k = \sum k_i$

  utilisation : si $X_i \sim \mathrm{N}(\mu, \sigma^2)$, alors $(n-1)\frac{S^2}{\sigma^2} \sim \chi^2_{n-1}$

- Student : *une loi* T_k *à* k *degrés de liberté a la distribution de* $T_k = \frac{Z}{\sqrt{\chi^2_k/k}}$ *où* $Z \sim N(0,1)$ *et* Z *et* χ^2_k *sont indépendantes*

  $X \sim T_k, \quad E(X) = 0, \quad Var(X) = \frac{k}{k-2}, \quad \text{quantile } P(T_k > t_{\alpha;k}) = \alpha$

  symétrie : $-t_{\alpha;k} = t_{1-\alpha;k}$

  utilisation : si $X_i \sim \mathrm{N}(\mu, \sigma^2)$, alors $\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim T_{n-1}$

Rappels Chapitre 8 - Échantillons Aléatoires et Lois d'Échantillonnage

- Fisher : *une loi $F_{u,v}$ à u et v degrés de liberté respectivement a la distribution de* $F_{u,v} = \frac{\chi^2_u/u}{\chi^2_v/v}$ *où χ^2_u et χ^2_v sont indépendantes*

  $X \sim F_{u,v}, \quad E(X) = \frac{v}{v-2}, \quad Var(X) = \frac{2v^2(u+v-2)}{u(v-2)^2(v-4)}, \quad P(F_{u,v} > f_{\alpha;u,v}) = \alpha$

  anti-symétrie : $f_{1-\alpha;u,v} = \frac{1}{f_{\alpha;u,v}}$

  utilisation : si $X_i \sim \mathrm{N}(\mu_X, \sigma^2_X)$, $Y_i \sim \mathrm{N}(\mu_Y, \sigma^2_Y)$, alors $\dfrac{S^2_X/\sigma^2_X}{S^2_Y/\sigma^2_Y} \sim F_{n_X-1,n_Y-1}$

Rappels Chapitre 9 - Estimation

- Biais d'un estimateur ponctuel $\hat{\theta}$: $Biais(\hat{\theta}) = E(\hat{\theta}) - \theta$
- Erreur quadratique moyenne de $\hat{\theta}$: $EQM(\hat{\theta}) = E((\hat{\theta} - \theta)^2) = Var(\hat{\theta}) + (Biais(\hat{\theta}))^2$

  $\Rightarrow \hat{\theta}_1$ meilleur que $\hat{\theta}_2$ si $EQM(\hat{\theta}_1) < EQM(\hat{\theta}^2)$
- Convergence d'un estimateur $\hat{\theta}$: si $\lim_{n\to\infty} EQM(\hat{\theta}) = 0$, alors $\hat{\theta}$ converge vers θ
- Méthode des moments : 1. calculer les $E(X^k) = h_k(\theta_1, \theta_2, ...)$,

  2. poser $h_k(\hat{\theta}_1, \hat{\theta}_2, ...) = \frac{1}{n}\sum_i X_i^k$,

  3. isoler les $\hat{\theta}_1$, $\hat{\theta}_2$, ...
- Maximum de vraisemblance : trouver les $\hat{\theta}_1$, $\hat{\theta}_2$, ... qui maximisent

  $L(\hat{\theta}_1, \hat{\theta}_2, ...) = \prod_{i=1}^{n} f_X(X_i, \hat{\theta}_1, \hat{\theta}_2, ...)$ (continu)

  $= \prod_{i=1}^{n} p_X(X_i, \hat{\theta}_1, \hat{\theta}_2, ...)$ (discret)

Rappels Chapitre 9 - Estimation

Situation	Loi utilisée	Intervalle de confiance: niveau $1-\alpha$
Une moyenne μ		
σ^2 est connue et $X \sim N(\mu, \sigma^2)$, ou n grand	$\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$\overline{X} - z_{\alpha/2}\cdot\dfrac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{\alpha/2}\cdot\dfrac{\sigma}{\sqrt{n}}$
σ^2 est inconnue et $X \sim N(\mu, \sigma^2)$	$\dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim T_{n-1}$	$\overline{X} - t_{\alpha/2;n-1}\cdot\dfrac{S}{\sqrt{n}} \leq \mu \leq \overline{X} + t_{\alpha/2;n-1}\cdot\dfrac{S}{\sqrt{n}}$
σ^2 est inconnue et n est très grand	$\dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim N(0,1)$	$\overline{X} - z_{\alpha/2}\cdot\dfrac{S}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{\alpha/2}\cdot\dfrac{S}{\sqrt{n}}$
Une variance σ^2		
$X \sim N(\mu, \sigma^2)$ et μ et σ^2 sont inconnues	$(n-1)\dfrac{S^2}{\sigma^2} \sim \chi^2_{n-1}$	$\dfrac{(n-1)S^2}{\chi^2_{\alpha/2;n-1}} \leq \sigma^2 \leq \dfrac{(n-1)S^2}{\chi^2_{1-\alpha/2;n-1}}$
Un écart-type σ		
(cas approximatif) n est très grand ($n \geq 40$)	$\dfrac{S-\sigma}{\sigma/\sqrt{2n}} \sim N(0,1)$	$\dfrac{S}{1+\dfrac{z_{\alpha/2}}{\sqrt{2n}}} \leq \sigma \leq \dfrac{S}{1-\dfrac{z_{\alpha/2}}{\sqrt{2n}}}$
Une proportion p		
$X \sim$ Bernoulli de paramètre p et n est très grand	$\dfrac{\hat{p}-p}{\sqrt{\dfrac{p(1-p)}{n}}} \sim N(0,1)$	$\hat{p} - z_{\alpha/2}\cdot\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + z_{\alpha/2}\cdot\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$

Attention z_α est tel que

$$\Phi(z_\alpha) = 1 - \alpha$$

Rappels Chapitre 9 - Estimation

Situation	Loi utilisée	Intervalle de confiance: niveau $1-\alpha$
Différence $\mu_1-\mu_2$		
σ_1^2 et σ_2^2 sont connues et $X_i \sim N(\mu_i,\sigma_i^2)$, $i=1,2$ ou alors n_1 et n_2 grands.	$\dfrac{\overline{X}_1-\overline{X}_2-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$	$(\overline{X}_1-\overline{X}_2) \pm z_{\alpha/2}\cdot\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}$
σ_1^2 et σ_2^2 sont inconnues avec $\sigma_1^2=\sigma_2^2=\sigma^2$ et $X_i \sim N(\mu_i,\sigma^2)$, $i=1,2$	$\dfrac{\overline{X}_1-\overline{X}_2-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim T_{n_1+n_2-2}$	$(\overline{X}_1-\overline{X}_2) \pm t_{\alpha/2;n_1+n_2-2}\cdot S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$ où $S_p=\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$
σ_1^2 et σ_2^2 sont inconnues avec $\sigma_1^2\neq\sigma_2^2$ et $X_i \sim N(\mu_i,\sigma_i^2)$, $i=1,2$	$\dfrac{\overline{X}_1-\overline{X}_2-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} \sim T_\nu$	$(\overline{X}_1-\overline{X}_2) \pm t_{\alpha/2;\nu}\cdot\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$ où $\nu=\frac{(S_1^2/n_1+S_2^2/n_2)^2}{\frac{(S_1^2/n_1)^2}{n_1+1}+\frac{(S_2^2/n_2)^2}{n_2+1}}-2$
σ_1^2 et σ_2^2 sont inconnues et n_1,n_2 sont grands ($n_1\geq 30$, $n_2\geq 30$)	$\dfrac{\overline{X}_1-\overline{X}_2-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} \sim N(0,1)$	$(\overline{X}_1-\overline{X}_2) \pm z_{\alpha/2}\cdot\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$
Les observations sont couplées $D=X_1-X_2\sim$ Normale $D_j=X_{1j}-X_{2j}$, $j=1,\dots,n$	$\dfrac{\overline{D}-(\mu_1-\mu_2)}{S_D/\sqrt{n}} \sim T_{n-1}$	$\overline{D}\pm t_{\alpha/2;n-1}\cdot\dfrac{S_D}{\sqrt{n}}$ où $\overline{D}=\frac{1}{n}\sum_{j=1}^{n}D_j, S_D=\sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(D_j-\overline{D})^2}$
Rapport σ_1^2/σ_2^2		
$X_i\sim N(\mu_i,\sigma_i^2)$, $i=1,2$	$\dfrac{S_2^2/\sigma_2^2}{S_1^2/\sigma_1^2}\sim F_{n_2-1;n_1-1}$	$L\leq\dfrac{\sigma_1^2}{\sigma_2^2}\leq U$ $L=\dfrac{S_1^2}{S_2^2}\cdot F_{1-\alpha/2;n_2-1,n_1-1}$ $U=\dfrac{S_1^2}{S_2^2}\cdot F_{\alpha/2;n_2-1,n_1-1}$
Différence p_1-p_2		
$X_1\sim$ Bernoulli $X_2\sim$ Bernoulli n_1 et n_2 sont grands	$\dfrac{\hat{p}_1-\hat{p}_2-(p_1-p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}}\sim N(0,1)$	$(\hat{p}_1-\hat{p}_2)\pm z_{\alpha/2}\cdot\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

Rappels Chapitre 10 - Tests d'Hypothèses

Situation	Statistique du test			
Une moyenne μ	$H_0 : \mu = \mu_0$	$H_1 : \mu < \mu_0$	$H_1 : \mu > \mu_0$	$H_1 : \mu \neq \mu_0$
σ^2 est connue et $X \sim N(\mu, \sigma^2)$, ou n grand	$Z_0 = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$	rejeter H_0 si $Z_0 < -z_\alpha$ $\beta(d)$, page 491	rejeter H_0 si $Z_0 > z_\alpha$ $\beta(d)$, page 491	rejeter H_0 si $\lvert Z_0 \rvert > z_{\alpha/2}$ $\beta(d)$, page 490
σ^2 est inconnue et $X \sim N(\mu, \sigma^2)$	$T_0 = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$	rejeter H_0 si $T_0 < -t_{\alpha;n-1}$ $\beta(d)$, page 493	rejeter H_0 si $T_0 > t_{\alpha;n-1}$ $\beta(d)$, page 493	rejeter H_0 si $\lvert T_0 \rvert > t_{\alpha/2;n-1}$ $\beta(d)$, page 492
σ^2 est inconnue et n est très grand	$Z_0 = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$	rejeter H_0 si $Z_0 < -z_\alpha$ $\beta(d)$, page 491	rejeter H_0 si $Z_0 > z_\alpha$ $\beta(d)$, page 491	rejeter H_0 si $\lvert Z_0 \rvert > z_{\alpha/2}$ $\beta(d)$, page 490
Une variance σ^2	$H_0 : \sigma^2 = \sigma_0^2$	$H_1 : \sigma^2 < \sigma_0^2$	$H_1 : \sigma^2 > \sigma_0^2$	$H_1 : \sigma^2 \neq \sigma_0^2$
$X \sim N(\mu, \sigma^2)$ et μ et σ^2 sont inconnues	$\chi_0^2 = (n-1)\dfrac{S^2}{\sigma_0^2}$	rejeter H_0 si $\chi_0^2 < \chi^2_{1-\alpha;n-1}$ $\beta(\lambda)$, page 496	rejeter H_0 si $\chi_0^2 > \chi^2_{\alpha;n-1}$ $\beta(\lambda)$, page 495	rejeter H_0 si $\chi_0^2 < \chi^2_{1-\alpha/2;n-1}$ ou $\chi_0^2 > \chi^2_{\alpha/2;n-1}$ $\beta(\lambda)$, page 494
n est grand ($n \geq 40$)	$Z_0 = \dfrac{S - \sigma_0}{\sigma_0/\sqrt{2n}}$	rejeter H_0 si $Z_0 < -z_\alpha$	rejeter H_0 si $Z_0 > z_\alpha$	rejeter H_0 si $\lvert Z_0 \rvert > z_{\alpha/2}$
Une proportion p	$H_0 : p = p_0$	$H_1 : p < p_0$	$H_1 : p > p_0$	$H_1 : p \neq p_0$
$X \sim$ Bernoulli de paramètre p et n est très grand	$Z_0 = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$	rejeter H_0 si $Z_0 < -z_\alpha$ si β : pp.291-292	rejeter H_0 si $Z_0 > z_\alpha$ β : pp.291-292	rejeter H_0 si $\lvert Z_0 \rvert > z_{\alpha/2}$ β : pp.291-292

p-value pour $H_0 : \mu = \mu_0$

σ^2 connue, unilatéral

$p\text{–}value = P(Z > |z_0|) = 1 - \Phi(|z_0|)$

σ^2 connue, bilatéral

$p - value = 2P(Z > |z_0|) = 2(1 - \Phi(|z_0|))$

σ^2 inconnue, unilatéral

$p - value = P(T > |z_0|)$ avec $k = n - 1$

σ^2 inconnue, bilatéral

$p - value = 2P(T > |z_0|)$ avec $k = n - 1$

$\alpha = P(\text{rejeter } H_0 | H_0 \text{ vraie})$

$\beta = P(\text{accepter } H_0 | H_0 \text{ fausse})$

Rappels Chapitre 10 - Tests d'Hypothèses

TABLE 11.2 – Calculs de β et de n (pour une moyenne avec **variance connue**)

Hypothèses	valeur de β	valeur de n
$H_0 : \mu = \mu_0$ contre $H_1 : \mu < \mu_0$	$\beta(\mu) = \Phi\left(z_\alpha + \frac{(\mu - \mu_0)\sqrt{n}}{\sigma}\right)$	$n = \frac{(z_\alpha + z_\beta)^2\sigma^2}{(\mu - \mu_0)^2}$
$H_0 : \mu = \mu_0$ contre $H_1 : \mu > \mu_0$	$\beta(\mu) = \Phi\left(z_\alpha - \frac{(\mu - \mu_0)\sqrt{n}}{\sigma}\right)$	$n = \frac{(z_\alpha + z_\beta)^2\sigma^2}{(\mu - \mu_0)^2}$
$H_0 : \mu = \mu_0$ contre $H_1 : \mu \neq \mu_0$	$\beta(\mu) = \Phi\left(z_{\alpha/2} - \frac{(\mu - \mu_0)\sqrt{n}}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \frac{(\mu - \mu_0)\sqrt{n}}{\sigma}\right)$	$n \approx \frac{(z_{\alpha/2} + z_\beta)^2\sigma^2}{(\mu - \mu_0)^2}$

$$\beta = P(\text{accepter } H_0 | H_0 \text{ fausse})$$

Rappels Chapitre 10 - Tests d'Hypothèses

Situation	Statistique du test			
2 moyennes μ_1, μ_2 | $H_0 : \mu_1 = \mu_2$		$H_1 : \mu_1 < \mu_2$	$H_1 : \mu_1 > \mu_2$	$H_1 : \mu_1 \neq \mu_2$
σ_1^2 et σ_2^2 sont connues et $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, 2$ ou alors n_1 et n_2 grands.	$Z_0 = \dfrac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	rejeter H_0 si $Z_0 < -z_\alpha$ $\beta(d)$, page 491	rejeter H_0 si $Z_0 > z_\alpha$ $\beta(d)$, page 491	rejeter H_0 si $\lvert Z_0 \rvert > z_{\alpha/2}$ $\beta(d)$, page 490
σ_1^2 et σ_2^2 sont inconnues avec $\sigma_1^2 = \sigma_2^2 = \sigma^2$ et $X_i \sim N(\mu_i, \sigma^2)$, $i = 1, 2$	$T_0 = \dfrac{\overline{X}_1 - \overline{X}_2}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ $(*)$	rejeter H_0 si $T_0 < -t_{\alpha;n_1+n_2-2}$	rejeter H_0 si $T_0 > t_{\alpha;n_1+n_2-2}$	rejeter H_0 si $\lvert T_0 \rvert > t_{\alpha/2;n_1+n_2-2}$
σ_1^2 et σ_2^2 sont inconnues avec $\sigma_1^2 \neq \sigma_2^2$ et $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, 2$	$T_0 = \dfrac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$ $(**)$	rejeter H_0 si $T_0 < -t_{\alpha;\nu}$	rejeter H_0 si $T_0 > t_{\alpha;\nu}$	rejeter H_0 si $\lvert T_0 \rvert > t_{\alpha/2;\nu}$
σ_1^2 et σ_2^2 sont inconnues et n_1, n_2 sont grands $(n_1 \geq 30,\ n_2 \geq 30)$	$Z_0 = \dfrac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$	rejeter H_0 si $Z_0 < -z_\alpha$	rejeter H_0 si $Z_0 > z_\alpha$	rejeter H_0 si $\lvert Z_0 \rvert > z_{\alpha/2}$
Les observations sont couplées $D = X_1 - X_2 \sim$ Normale $D_j = X_{1j} - X_{2j}$, $j = 1, \ldots, n$	$T_0 = \dfrac{\overline{D}}{S_D/\sqrt{n}}$ $(***)$	rejeter H_0 si $T_0 < -t_{\alpha;n-1}$	rejeter H_0 si $T_0 > t_{\alpha;n-1}$	rejeter H_0 si $\lvert T_0 \rvert > t_{\alpha/2;n-1}$
2 variances σ_1^2, σ_2^2 | $H_0 : \sigma_1^2 = \sigma_2^2$		$H_1 : \sigma_1^2 < \sigma_2^2$	$H_1 : \sigma_1^2 > \sigma_2^2$	$H_1 : \sigma_1^2 \neq \sigma_2^2$
$X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, 2$	$F_0 = \dfrac{S_1^2}{S_2^2}$	rejeter H_0 si $F_0 < F_{1-\alpha;n_1-1,n_2-1}$	rejeter H_0 si $F_0 > F_{\alpha;n_1-1,n_2-1}$ $\beta(\lambda)$, page 498	rejeter H_0 si $F_0 < F_{1-\alpha/2;n_1-1,n_2-1}$ ou $F_0 > F_{\alpha/2;n_1-1,n_2-1}$ $\beta(\lambda)$, page 497
n_1 et n_2 sont grands $n_1 > 40,\ n_2 > 40$	$Z_0 = \dfrac{S_1 - S_2}{S_p\sqrt{\frac{1}{2n_1} + \frac{1}{2n_2}}}$ $(*)$	rejeter H_0 si $Z_0 < -z_\alpha$	rejeter H_0 si $Z_0 > z_\alpha$	rejeter H_0 si $\lvert Z_0 \rvert > z_{\alpha/2}$

$(*)\ S_p = \sqrt{\dfrac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}$ $\quad (**)\ \nu = \dfrac{(S_1^2/n_1 + S_2^2/n_2)^2}{\frac{(S_1^2/n_1)^2}{n_1+1} + \frac{(S_2^2/n_2)^2}{n_2+1}} - 2$ $\quad (***)\ \overline{D} = \dfrac{1}{n}\sum_{j=1}^{n} D_j,\ S_D = \sqrt{\dfrac{1}{n-1}\sum_{j=1}^{n}(D_j}$

Rappels Chapitre 10 - Tests d'Hypothèses

TABLE 11.4 – Calculs de β et de n (pour deux moyennes avec **variances connues**)

Hypothèses	valeur de β	valeur de n (on suppose $n = n_1 = n_2$)
$H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 < \mu_2$	$\beta(\mu_1, \mu_2) = \Phi\left(z_\alpha + \dfrac{(\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \right)$	$n = \dfrac{(z_\alpha + z_\beta)^2(\sigma_1^2 + \sigma_2^2)}{(\mu_1 - \mu_2)^2}$
$H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 > \mu_2$	$\beta(\mu_1, \mu_2) = \Phi\left(z_\alpha - \dfrac{(\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \right)$	$n = \dfrac{(z_\alpha + z_\beta)^2(\sigma_1^2 + \sigma_2^2)}{(\mu_1 - \mu_2)^2}$
$H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 \neq \mu_2$	$\beta(\mu_1, \mu_2) = \Phi\left(z_{\alpha/2} - \dfrac{(\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \right) - \Phi\left(-z_{\alpha/2} - \dfrac{(\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \right)$	$n \approx \dfrac{(z_{\alpha/2} + z_\beta)^2(\sigma_1^2 + \sigma_2^2)}{(\mu_1 - \mu_2)^2}$

- Pour tester $H_0 : X \sim \text{Loi}(\theta_1, ...)$ contre $H_1 : X \not\sim \text{Loi}(\theta_1, ...)$ dont p paramètres sont estimés (non connus), étant donné

Valeurs (x_i)	V_1	V_2	...	V_k	Total
Effectifs observés (O_i)	O_1	O_2	...	O_k	n
Effectifs attendus (E_i)	E_1	E_2	...	E_k	n

où $E_i = n \cdot P(X \in V_i | H_0 \text{ est vraie})$,

on considère $\chi_0^2 = \displaystyle\sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$ et on rejette H_0 si $\chi_0^2 > \chi^2_{\alpha;k-p-1}$

Rappels Chapitre 10 - Tests d'Hypothèses

- Pour tester H_0 : X et Y indép. contre H_1 : X et Y non indép., étant donné

X \ Y	y_1	y_2	$\cdots$	y_j	$\cdots$	y_c	Total
x_1	O_{11}	O_{12}	$\cdots$	O_{1j}	$\cdots$	O_{1c}	$\sum_{j=1}^{c} O_{1j}$
x_2	O_{21}	O_{22}	$\cdots$	O_{2j}	$\cdots$	O_{2c}	$\sum_{j=1}^{c} O_{2j}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_i	O_{i1}	O_{i2}	$\cdots$	O_{ij}	$\cdots$	O_{ic}	$\sum_{j=1}^{c} O_{ij}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_r	O_{r1}	O_{r2}	$\cdots$	O_{rj}	$\cdots$	O_{rc}	$\sum_{j=1}^{c} O_{rj}$
Total	$\sum_{i=1}^{r} O_{i1}$	$\sum_{i=1}^{r} O_{i2}$	$\cdots$	$\sum_{i=1}^{r} O_{ij}$	$\cdots$	$\sum_{i=1}^{r} O_{ic}$	n

on considère $\chi_0^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$ où $E_{ij} = \frac{1}{n} \left(\sum_{k=1}^{c} O_{ik} \right) \left(\sum_{l=1}^{r} O_{lj} \right)$

on rejette H_0 si $\chi_0^2 > \chi^2_{\alpha;(r-1)(c-1)}$

Rappels Chapitre 11 - Régression Linéaire

I-Calcul des estimateurs

Cas particulier : Régression simple

$$\hat{\beta}_1 = \frac{S_{XY}}{S_{XX}} \text{ et } \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

$$S_{XX} = \sum (X_i - \bar{X})^2 = \sum X_i^2 - n\bar{X}^2$$

$$S_{YY} = \sum (Y_i - \bar{Y})^2 = \sum Y_i^2 - n\bar{Y}^2$$

$$S_{XY} = \sum (X_i - \bar{X})(Y_i - \bar{Y}) = \sum X_i Y_i - n\bar{X}\bar{Y}$$

II-Test de signification

a-Test global (loi de Fisher)

$H_0: \beta_1 = 0 \;\; vs \;\; H_1: \beta_1 \neq 0.$

L'hypothèse $\mathbf{H_0 : \beta_1 = 0}$ est **rejetée**, au seuil α, si $\mathbf{F_0 > F_{1,n-2}(\alpha)}$.
Lorsque H_0 est **rejetée**, le modèle est **globalement significatif**, donc la variable X est **significative**.
On peut aussi effectuer un test de signification individuel pour la variable X. C'est un test équivalent au test global pour une régression simple.

Tableau d'analyse de variance – Régression simple

Source de variation	Somme des carrés (khi-2)	Deg liberté	Moyenne des carrés	F (statistique)	p-value
Régression (modèle)	$SS_R = \sum(\hat{y}_i - \bar{y})^2 = \hat{\beta}_1 S_{XY}$	1	$MS_R = SS_R / 1$	$F_0 = MS_R / MS_E$	$P(F \geq F_0)$
Résidus (Erreur)	$SS_E = \sum(y_i - \hat{y}_i)^2 = S_{YY} - SS_R$	n-2	$MS_E = SS_E / n-2$		
Totale	$S_{YY} = SC_T = \sum(y_i - \bar{y})^2$	n-1			

Rappels Chapitre 11 - Régression Linéaire

$$R^2 = SS_R / S_{YY} = 1 - SS_E / S_{YY} \qquad R^2_{ajusté} = \frac{(n-1)R^2 - 1}{n-2}$$

$$\hat{\sigma}^2 = MS_E$$

$$\beta_1 \in \hat{\beta}_1 \pm t_{\alpha/2;n-2}\sqrt{\frac{MS_E}{S_{xx}}} \qquad \beta_0 \in \hat{\beta}_0 \pm t_{\alpha/2;n-2}\sqrt{MS_E\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)}$$

<u>Confiance</u> : $E(Y|X = x_0) \in \hat{y}_0 \mp t_{n-2}(\alpha/2)\sqrt{MS_E(\frac{1}{n} + \frac{(\bar{X}-x_0)^2}{S_{XX}})}$

<u>Prévison</u> : $(Y|X = x_0) \in \hat{y}_0 \mp t_{n-2}(\alpha/2)\sqrt{MS_E(1 + \frac{1}{n} + \frac{(\bar{X}-x_0)^2}{S_{XX}})}$

Rappels Chapitre 12 - Fiabilité

- Fiabilité : $R(t) = P(T > t) = 1 - F_T(t)$
- Durée de vie moyenne : $\tau = E(T) = \int_0^\infty R(t)\mathrm{d}t$
- Taux de panne : $r(t) = -\frac{R'(t)}{R(t)}$, tel que $R(t) = \exp\left(-\int_0^t r(x)\mathrm{d}x\right)$
- Si $T \sim \mathrm{Exp}(\lambda)$, alors $R(t) = e^{-\lambda t}$, $\tau = \frac{1}{\lambda}$

- En série : $T = \min\{T_1, ..., T_n\}$, $R(t) = \prod_{i=1}^n R_i(t)$
- En paral., redondance active : $T = \max\{T_1, ..., T_n\}$, $R(t) = 1 - \prod_{i=1}^n (1 - R_i(t))$
- En paral., redondance passive : $T = T_1 + ... + T_n$, $\tau = \tau_1 + ... + \tau_n$

Rappels Chapitre 13 - Files d'Attente

- Arrivées clients suit processus de Poisson de taux λ $\rho = \frac{\lambda}{\mu}$
- Temps de service $\sim \text{Exp}(\mu)$
- Nombre total (attente + service) de clients en file à l'éq. N avec $N+1 \sim \text{G}(1-\rho)$
- Temps total dans la file (attente + service) à l'équilibre $T \sim \text{Exp}(\mu - \lambda)$

- Nombre total (attente + service) moyen de clients en file à l'équilibre : $\bar{N} = \frac{\rho}{1-\rho}$
- Nombre moyen de clients en service à l'équilibre : $\bar{N}_S = \rho$
- Nombre moyen de clients en attente à l'équilibre : $\bar{N}_Q = \bar{N} - \bar{N}_S = \frac{\rho^2}{1-\rho}$
- Temps total dans la file (attente + service) moyen à l'équilibre : $\bar{T} = \frac{\bar{N}}{\lambda} = \frac{1}{\mu-\lambda}$
- Temps en service moyen à l'équilibre : $\bar{T}_S = \frac{1}{\mu}$
- Temps en attente moyen à l'équilibre : $\bar{T}_Q = \bar{T} - \bar{T}_S = \frac{\lambda}{\mu(\mu-\lambda)}$