

MATH-H204 - Calcul des probabilités et statistiques

Yves DE SMET

Résumé du cours

Rodrigue VAN BRANDE

Dylan GONZALEZ

29 août 2015

Table des matières

1	La statistique descriptive	4
1.1	Statistique descriptive en 1D	4
1.2	Statistique descriptive en 2D	4
1.2.1	Covariance	4
1.2.2	Le coefficient de corrélation	5
1.2.3	Les droites de régression	6
1.2.4	Variances résiduelles	8
2	La théorie des probabilités	9
2.1	Probabilités	9
2.1.1	Axiomes de la théorie des probabilités	9
2.1.2	Probabilité indépendance	9
2.1.3	Probabilité conditionnelle et indépendance	9
2.1.4	Probabilité conditionnelle inverse et indépendant	9
2.1.5	Formule de Bayes	10
2.2	Variables aléatoires	11
2.2.1	Valeurs typiques	11
2.2.2	Distribution d'une fonction monotone d'une variable aléatoire $G(V)$	12
2.2.3	Distribution de la somme de deux variables aléatoires $V + W$	13
2.2.4	Distribution du produit de deux variables aléatoires $V.W$	14
2.2.5	Distribution de la somme des espérances $E(V + W)$	15
2.2.6	Distribution du produit des espérances $E(V.W)$	16
2.2.7	Variance d'une variable aléatoire $D^2(aV + b)$	17
2.2.8	Variance de la somme de deux variables aléatoires $D^2(V + W)$	17
2.3	Variables aléatoires particulières	18
2.3.1	Schéma de relation entre les variables aléatoires	18
2.3.2	Variable binomiale $\mathcal{B}(n, p)$	18
2.3.3	Variable de Poisson \mathcal{P}_λ	19
2.3.4	Variable exponentielle négative	21
2.3.5	Variable Normale $\mathcal{N}(\mu, \sigma)$	21
2.3.6	Variable Khi^2	21
2.3.7	Variable Student t_n	21
2.3.8	Variable Snedecor $\mathcal{F}_{(m,n)}$	21
2.4	Théorèmes fondamentaux	22
2.4.1	Inégalité de Bienaymé-Tchebycheff	22
2.4.2	Théorème de Bernoulli ou loi des grands nombres	23
2.4.3	Théorème Central-Limite	24
2.4.4	Théorème de De Moivre	27
3	L'inférence statistique	28
3.1	Les valeurs	28
3.2	Distributions échantillonnées	28
3.2.1	Distribution échantillonnées de la moyenne \bar{X}	28
3.2.2	Distribution échantillonnées de la variance S^2	28
3.2.3	Distribution échantillonnées d'une fonction de \bar{X} et s^2	29
3.3	Test d'hypothèse	30
3.4	Intervalles de confiance et tests d'hypothèses	31
3.4.1	Comparaison des moyennes de deux populations normales de deux écart-type σ connu	31
3.4.2	Comparaison des moyennes de deux populations normales de même écart-type σ inconnu	32
3.4.3	Comparaison des moyennes de deux populations quelconques	33

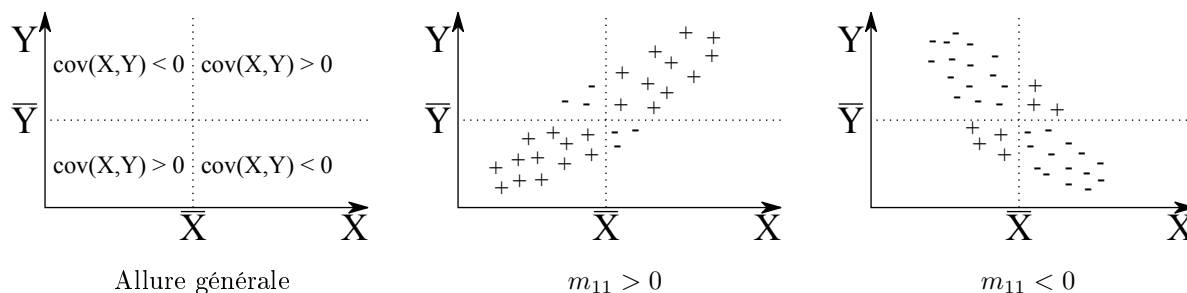
4	Autres	34
4.1	Tableau du formulaire	34
4.2	Densité et répartition	34
4.3	Distributions	35

1 La statistique descriptive

1.1 Statistique descriptive en 1D

1.2 Statistique descriptive en 2D

1.2.1 Covariance



La covariance $|m_{11}| \leq s_1 s_2$

La covariance est le moment d'ordre (1,1) :

$$\begin{aligned}
 m_{11} &= \frac{1}{n} \sum_{i=1}^p \sum_{j=1}^q n_{ij} (x_i - \bar{x})(y_j - \bar{y}) \\
 0 &\leq \frac{1}{n} \sum_{i=1}^p \sum_{j=1}^q n_{ij} \underbrace{(u(x_i - \bar{x}) + (y_j - \bar{y}))^2}_{\substack{2 \text{ car toujours } \geq 0}} \\
 &\leq \frac{1}{n} \sum_{i=1}^p \sum_{j=1}^q n_{ij} (u^2(x_i - \bar{x})^2 + 2u(x_i - \bar{x})(y_j - \bar{y}) + (y_j - \bar{y})^2) \\
 &\leq u^2 s_1^2 + 2u m_{11} + s_2^2
 \end{aligned}$$

Équation du second degré, on calcule son Δ :

$$\begin{aligned}
 \Delta &\leq 0 \\
 (2m_{11})^2 - 4s_1^2 s_2^2 &\leq 0 \\
 m_{11}^2 - s_1^2 s_2^2 &\leq 0 \\
 m_{11}^2 &\leq s_1^2 s_2^2
 \end{aligned}$$

$$\boxed{|m_{11}| \leq s_1 s_2}$$

La covariance maximale $|m_{11}| = s_1 s_2$

La valeur absolue de la covariance est maximale si elle vaut $|m_{11}| = s_1 s_2$.

Si les points observés se trouvent sur une droite non parallèle aux axes $ax + by + c = 0$, on a $ax_i + by_i + c = 0$.

On multiplie par $\frac{n_{ij}}{n}$ et on somme sur ij .

$$\begin{aligned}
 0 &= \sum_{i=1}^p \sum_{j=1}^q \frac{n_{ij}}{n} (ax_i + by_j + c) \\
 &= a \frac{1}{n} \sum_{i=1}^p \sum_{j=1}^q n_{ij} x_i + b \frac{1}{n} \sum_{i=1}^p \sum_{j=1}^q n_{ij} y_j + c \frac{1}{n} \sum_{i=1}^p \sum_{j=1}^q n_{ij} \\
 &= a\bar{x} + b\bar{y} + c
 \end{aligned}$$

On soustrait $ax_i + by_j + c = 0$ par $a\bar{x} + b\bar{y} + c = 0$.

$$= a(x_i - \bar{x}) + b(y_j - \bar{y})$$

On utilise $u_0 = \frac{a}{b}$

$$= u_0 b(x_i - \bar{x}) + b \frac{a}{u_0} (y_j - \bar{y})$$

$$= u_0 (x_i - \bar{x}) + (y_j - \bar{y})$$

L'équation a la même forme que α , du coup...

$$0 = \Delta$$

$$= m_{11}^2 - s_1^2 s_2^2$$

$$m_{11}^2 = s_1^2 s_2^2$$

$$\boxed{|m_{11}| = s_1 s_2}$$

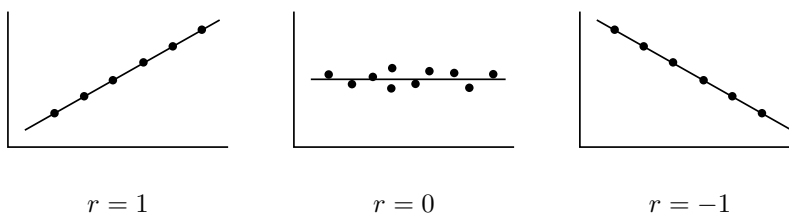
1.2.2 Le coefficient de corrélation

$$\boxed{r = \frac{m_{11}}{s_1 s_2}}$$

Propriétés

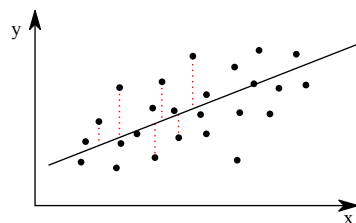
1. r sans dimensions ;
2. $r' = r$;
3. $-1 \leq r \leq 1$;
4. $|r| = 1$ ssi les points observés se trouvent sur une droite non parallèle aux axes.

Représentation de la droite de régression en fonction du coefficient de corrélation



- si $|r| = 1$, il y a une relation fonctionnelle linéaire entre X et Y ;
- si $r = 0$, Y est indépendante de X : la covariance est nulle et la droite de régression est horizontale ;
- la liaison entre X et Y est d'autant plus intime que $|r|$ est voisin de 1, et d'autant plus faible que $|r|$ est voisin de 0.

1.2.3 Les droites de régression



La droite de régression de y en x est la droite qui minimise la somme des carrés des écarts des points observés à cette droite, les écarts étant pris parallèlement à l'axe y . C'est donc la droite d'équation $y = ax + b$ qui minimise la quantité.

$$g(a, b) = \sum_{i=1}^p \sum_{j=1}^q n_{ij} (y_j - a x_i - b)^2$$

Dérivée par rapport à a :

$$\begin{aligned}
 0 &= g(a, b)|_a \\
 \Leftrightarrow 0 &= \sum_{i=1}^p \sum_{j=1}^q n_{ij} 2(y_j - a x_i - b)(-x_i) \\
 \Leftrightarrow 0 &= \sum_{i=1}^p \sum_{j=1}^q -2n_{ij} x_i (y_j - a x_i - b) \\
 \Leftrightarrow 0 &= \sum_{i=1}^p \sum_{j=1}^q n_{ij} x_i (-y_j + a x_i + b) \\
 \Leftrightarrow 0 &= \sum_{i=1}^p \sum_{j=1}^q -n_{ij} x_i y_j + \sum_{i=1}^p \sum_{j=1}^q n_{ij} a x_i^2 + \sum_{i=1}^p \sum_{j=1}^q n_{ij} b x_i \\
 \Leftrightarrow 0 &= -1 \sum_{i=1}^p \sum_{j=1}^q n_{ij} x_i y_j + a \sum_{i=1}^p \sum_{j=1}^q n_{ij} x_i^2 + b \sum_{i=1}^p \sum_{j=1}^q n_{ij} x_i \\
 \Leftrightarrow \sum_{i=1}^p \sum_{j=1}^q n_{ij} x_i y_j &= \underbrace{a \sum_{i=1}^p \sum_{j=1}^q n_{ij} x_i^2 + b \sum_{i=1}^p \sum_{j=1}^q n_{ij} x_i}_{\substack{\text{Il n'existe aucun terme en } j \\ \text{Donc on peut supprimer la somme } \sum_{j=1}^q}} \\
 \Leftrightarrow \sum_{i=1}^p \sum_{j=1}^q n_{ij} x_i y_j &= \overbrace{a \sum_{i=1}^p n_{i.} x_i^2 + b \sum_{i=1}^p n_{i.} x_i}
 \end{aligned}$$

$$\sum_{i=1}^p \sum_{j=1}^q n_{ij} x_i y_j = a \sum_{i=1}^p n_{i.} x_i^2 + b \sum_{i=1}^p n_{i.} x_i$$

Dérivée par rapport à b :

$$\begin{aligned}
0 &= g(a, b)|_b \\
\Leftrightarrow 0 &= \sum_{i=1}^p \sum_{j=1}^q n_{ij} 2(y_j - a x_i - b)(-1) \\
\Leftrightarrow 0 &= \sum_{i=1}^p \sum_{j=1}^q -2n_{ij}(y_j - a x_i - b) \\
\Leftrightarrow 0 &= \sum_{i=1}^p \sum_{j=1}^q n_{ij}(-y_j + a x_i + b) \\
\Leftrightarrow 0 &= \sum_{i=1}^p \sum_{j=1}^q -n_{ij} y_j + \sum_{i=1}^p \sum_{j=1}^q n_{ij} a x_i + \sum_{i=1}^p \sum_{j=1}^q n_{ij} b \\
\Leftrightarrow 0 &= -1 \sum_{i=1}^p \sum_{j=1}^q n_{ij} y_j + a \sum_{i=1}^p \sum_{j=1}^q n_{ij} x_i + b \sum_{i=1}^p \sum_{j=1}^q n_{ij} \\
\Leftrightarrow \sum_{i=1}^p \sum_{j=1}^q n_{ij} y_j &= a \sum_{i=1}^p \sum_{j=1}^q n_{ij} x_i + b \sum_{i=1}^p \sum_{j=1}^q n_{ij} \\
\Leftrightarrow n \bar{y} &= a n \bar{x} + b n
\end{aligned}$$

$$\boxed{n \bar{y} = a n \bar{x} + b n}$$

On a obtenu ces deux réponses :

$$\begin{cases} \sum_{i=1}^p \sum_{j=1}^q n_{ij} x_i y_j = a \sum_{i=1}^p n_{i.} x_i^2 + b \sum_{i=1}^p n_{i.} x_i & (1) \\ n \bar{y} = a n \bar{x} + b n & (2) \end{cases}$$

On soustrait le (1) par le double du (2) :

$$\begin{aligned}
&\bar{x} \cdot (2) : & n \bar{y} \bar{x} &= a n \bar{x}^2 + b n \bar{x} \\
(1) - \bar{x} \cdot (2) : & \left(\sum_{i=1}^p \sum_{j=1}^q n_{ij} x_i y_j \right) - (n \bar{y} \bar{x}) &= \left(a \sum_{i=1}^p n_{i.} x_i^2 + b \sum_{i=1}^p n_{i.} x_i \right) - (a n \bar{x}^2 + b n \bar{x}) \\
& & \vdots &
\end{aligned}$$

Pour obtenir à la fin :

$$\boxed{a = \frac{m_{11}}{s_1^2}} \text{ et } \boxed{b = \bar{y} - \frac{m_{11}}{s_1^2} \bar{x}}$$

On remplace dans une droite :

$$\begin{aligned}
y &= ax + b \\
y &= \frac{m_{11}}{s_1^2} x + \bar{y} - \frac{m_{11}}{s_1^2} \bar{x}
\end{aligned}$$

Pour obtenir la droite de regression de y en x :

$$\boxed{y = \frac{m_{11}}{s_1^2} (x - \bar{x}) + \bar{y}}$$

Le raisonnement est symétrique pour le cas de la régression de x en y :

$$\boxed{x = \frac{m_{11}}{s_2^2} (y - \bar{y}) + \bar{x}}$$

1.2.4 Variances résiduelles

La variance résiduelle de y en x est :

$$s_{21}^2 = s_2^2(1 - r^2)$$

Propriété et interprétation

- $s_{21}^2 = 0$ ssi $r = \pm 1$ ssi tous les points observés sont sur une droite ;
- $s_{21}^2 = s_2^2$ ssi $r = 0$ ssi les droites de régression sont parallèles aux axes ;
- $s_2^2 r^2$ représente une certaine proportion de s_2^2 , d'autant plus grande que la dépendance linéaire entre x et y est forte : on peut donc la considérer comme la part de la variance de y qui est expliquée par le lien linéaire entre x et y , tandis que la variance résiduelle s_{21}^2 est la part de la variance de y qui n'est pas expliquée par ce lien linéaire (d'où son nom).

Démonstration

Nous partons de la variance de y en x :

$$\begin{aligned}
 s_{21}^2 &= \frac{1}{n} \sum_{i=1}^p \sum_{j=1}^q n_{ij} (y_i - ax_i - b)^2 \\
 &\quad \underbrace{\hspace{10em}}_{\text{valeur minimum de } g(a,b)} \\
 &\quad \boxed{a = \frac{m_{11}}{s_1^2}} \text{ et } \boxed{b = \bar{y} - \frac{m_{11}}{s_1^2} \bar{x}} \\
 &= \frac{1}{n} \sum_{i=1}^p \sum_{j=1}^q n_{ij} \left(y_i - \frac{m_{11}}{s_1^2} x_i - \bar{y} + \frac{m_{11}}{s_1^2} \bar{x} \right)^2 \\
 &= \frac{1}{n} \sum_{i=1}^p \sum_{j=1}^q n_{ij} \left((y_i - \bar{y}) - \frac{m_{11}}{s_1^2} (x_i - \bar{x}) \right)^2 \\
 &= \frac{1}{n} \sum_{i=1}^p \sum_{j=1}^q n_{ij} (y_i - \bar{y})^2 - 2 \frac{m_{11}}{s_1^2} (x_i - \bar{x})(y_i - \bar{y}) + \frac{m_{11}^2}{s_1^4} (x_i - \bar{x})^2 \\
 &= \frac{1}{n} \sum_{i=1}^p \sum_{j=1}^q n_{ij} (y_i - \bar{y})^2 - 2 \frac{m_{11}}{s_1^2} \frac{1}{n} \sum_{i=1}^p \sum_{j=1}^q n_{ij} (x_i - \bar{x})(y_i - \bar{y}) + \frac{m_{11}^2}{s_1^4} \frac{1}{n} \sum_{i=1}^p \sum_{j=1}^q n_{ij} (x_i - \bar{x})^2 \\
 &= s_2^2 - 2 \frac{m_{11}}{s_1^2} m_{11} + \frac{m_{11}^2}{s_1^4} s_1^2 \\
 &= s_2^2 - 2 \frac{m_{11}^2}{s_1^2} + \frac{m_{11}^2}{s_1^2} \\
 &= s_2^2 - \frac{m_{11}^2}{s_1^2}
 \end{aligned}$$

On utilise le coefficient de corrélation $r = \frac{m_{11}}{s_1 s_2}$, donc $\frac{m_{11}}{s_1} = r s_2$:

$$\begin{aligned}
 s_2^2 - \frac{m_{11}^2}{s_1^2} &= s_2^2 - (r s_2)^2 \\
 &= s_2^2(1 - r^2)
 \end{aligned}$$

$$s_{21}^2 = s_2^2(1 - r^2)$$

La variance résiduelle de x en y est similaire :

$$s_{12}^2 = s_1^2(1 - r^2)$$

2 La théorie des probabilités

2.1 Probabilités

2.1.1 Axiomes de la théorie des probabilités

$$\left\{ \begin{array}{l} P(E) = 1 \\ P(A) \geq 0 \\ A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B) \end{array} \right. \quad \forall A \subset E$$

2.1.2 Probabilité indépendance

Si A et B sont indépendants alors

- A et \bar{B} sont indépendants ;
- \bar{A} et \bar{B} sont indépendants ;
- \bar{A} et B sont indépendants.

2.1.3 Probabilité conditionnelle et indépendance

Probabilité conditionnelle de A sous la condition B ("sachant B ") :

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

$$\text{Si } A \text{ est indépendant de } B : \left\{ \begin{array}{l} P(A \cap B) = P(A)P(B) \\ P(A|B) = P(A) \\ P(B|A) = P(B) \end{array} \right.$$

Une des propriétés des 3 suffit.

Démonstration

La probabilité conditionnelle respecte les axiomes de la théorie des probabilités

$$\begin{aligned} P(E|A) &= \frac{P(E \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1 \\ P(B|A) &= \frac{P(B \cap A)}{P(A)} \geq 0 \\ X \cap Y \neq \emptyset &\Rightarrow \frac{P(X \cup Y|A)}{\frac{P((X \cup Y) \cap A)}{P(A)}} \\ &= \frac{P((X \cap A) \cup (Y \cap A))}{\frac{P(X \cap A) + P(Y \cap A)}{P(A)}} \\ &= \frac{P(X|A)P(A) + P(Y|A)P(A)}{P(A)} \\ &= P(X|A) + P(Y|A) \end{aligned}$$

2.1.4 Probabilité conditionnelle inverse et indépendant

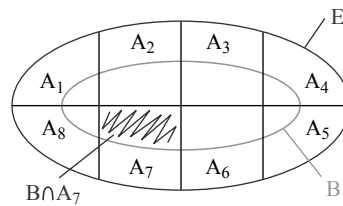
$$P(A|B) = P(A|\bar{B})$$

Démonstration

$$\begin{aligned}
P(A|B) &= P(A|\bar{B}) \\
\frac{P(A \cap B)}{P(B)} &= \frac{P(A \cap \bar{B})}{P(\bar{B})} \\
\frac{P(A)P(B)}{P(B)} &= \frac{P(A)P(\bar{B})}{P(\bar{B})} \\
P(A) &= P(A)
\end{aligned}$$

2.1.5 Formule de Bayes

Cette formule est utilisée dans le cas où un événement B peut survenir à cause d'évènement A_i incompatibles. Par exemple : une pièce défectueuse fabriquée par plusieurs machines différentes.



$$\begin{aligned}
B &= (A_1 \cap B) \cup (A_2 \cap B) \cup \dots \cup (A_m \cap B) \\
P(B) &= P(A_1 \cap B) + P(A_2 \cap B) + \dots + P(A_m \cap B) \\
&= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \dots + P(B|A_m)P(A_m)
\end{aligned}$$

$$P(A_k|B) = \frac{A_k \cap B}{P(B)} = \frac{P(B|A_k)P(A_k)}{\sum_{j=1}^m P(B|A_j)P(A_j)}$$

2.2 Variables aléatoires

2.2.1 Valeurs typiques

Aux distributions/densités marginales et conditionnelles, on peut associer les valeurs typiques :

Moyenne marginale μ_1	$= \sum_i p_{i.} x_i$	$= \int_{-\infty}^{\infty} x f_1$
Variance marginale σ_1^2	$= \sum_i p_{i.} (x_i - \mu_1)^2$	$= \int_{-\infty}^{\infty} (x - \mu_1)^2 f_1$
Moyenne conditionnelle	$\mu_{1/j} = \sum_i p_{i/j} x_i$	$\mu_{1/y_0} = \int_{-\infty}^{\infty} x f(x - \mu_1)^2 f_1$
Variance conditionnelle	$\sigma_{1/j}^2 = \sum_i p_{i/j} (x_i - \mu_{1/j})^2$	$\sigma_{1/y_0}^2 = \int_{-\infty}^{\infty} (x - \mu_{1/y_0})^2 f(x/y_0)$

2.2.2 Distribution d'une fonction monotone d'une variable aléatoire $G(V)$

$$\boxed{W = G(V)}$$

Cas discret

$$\begin{aligned} F_W(x) &= P(W \leq x) \\ &= P(G(V) \leq x) \\ &= P(V \leq G^{-1}(x)) \\ &= F_V(G^{-1}(x)) \end{aligned}$$

Cas continu

$$\begin{aligned} f_W(x) &= \begin{array}{c|c} \text{croissante} & \text{décroissante} \end{array} \\ &= F'_W(x) \\ &= (F_W(x))' \\ &= (P(W \leq x))' \\ &= (P(G(V) \leq x))' \\ &= \begin{array}{c|c} (P(V \leq G^{-1}(x)))' & (P(V \geq G^{-1}(x)))' \\ (F_V(G^{-1}(x)))' & (1 - F_V(G^{-1}(x)))' \\ f_V(G^{-1}(x)) (G^{-1}(x))' & -f_V(G^{-1}(x)) (G^{-1}(x))' \end{array} \\ \text{Rappel : } &\left\{ \begin{array}{l|l} G(G^{-1}(x)) & = x \\ (G(G^{-1}(x)))' & = (x)' \\ G'(G^{-1}(x)) (G^{-1}(x))' & = 1 \\ (G^{-1}(x))' & = \frac{1}{G'(G^{-1}(x))} \end{array} \right. \\ &= f_V(G^{-1}(x)) \frac{1}{G'(G^{-1}(x))} \left| f_V(G^{-1}(x)) \frac{-1}{G'(G^{-1}(x))} \right. \end{aligned}$$

2.2.3 Distribution de la somme de deux variables aléatoires $V + W$

$$\boxed{Z = V + W}$$

Cas discret

$$\begin{aligned} F_Z(x) &= \sum_i \sum_j p_{ij} \\ &= \sum_i \sum_j P(V \leq v_i, W \leq w_i) \quad \text{tel que } v_i + w_i \leq x \end{aligned}$$

Cas continu

$$\begin{aligned} F_Z(x) &= P(Z \leq x) \\ F_{V+W}(x) &= P(V \leq \xi, W \leq \eta) \quad \text{tel que } v_i + w_i \leq x \\ &= \iint_{\xi+\eta \leq x} f_{(V,W)}(\xi, \eta) \, d\xi \, d\eta \\ \text{On remplace par } &\left\{ \begin{array}{l} \xi = u \\ \eta = v - u \\ J = \begin{pmatrix} \frac{\partial \xi}{\partial u} & \frac{\partial \xi}{\partial v} \\ \frac{\partial \eta}{\partial u} & \frac{\partial \eta}{\partial v} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = 1.1 - 0.(-1) = 1 \end{array} \right. \\ &= \iint_{v \leq x} f_{(V,W)}(u, v - u) |1| \, du \, dv \\ &= \int_{-\infty}^x dv \int_{-\infty}^{+\infty} f_{(V,W)}(u, v - u) \, du \end{aligned}$$

$$\boxed{f_Z(x) = \frac{\delta F_Z(x)}{\delta x} = \int_{-\infty}^{+\infty} f_{(V,W)}(u, x - u) \, du}$$

ou si indépendant

$$\boxed{f_Z(x) = \frac{\delta F_Z(x)}{\delta x} = \int_{-\infty}^{+\infty} f_V(u) \cdot f_W(x - u) \, du}$$

2.2.4 Distribution du produit de deux variables aléatoires $V.W$

$$\boxed{Z = V.W}$$

Cas continu

$$\begin{aligned}
 F_Z(x) &= P(Z \leq x) \\
 F_{V.W}(x) &= P(V \leq \xi, W \leq \eta) \quad \text{tel que } v_i + w_i \leq x \\
 &= \iint_{\xi, \eta \leq x} f_{(V,W)}(\xi, \eta) d\xi d\eta \\
 \text{On remplace par } &\begin{cases} \xi = u \\ \eta = \frac{v}{u} \\ J = \begin{pmatrix} \frac{\partial \xi}{\partial u} & \frac{\partial \xi}{\partial v} \\ \frac{\partial \eta}{\partial u} & \frac{\partial \eta}{\partial v} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{u^2} & \frac{1}{u} \end{pmatrix} = 1 \cdot \frac{1}{u} - 0 \cdot \frac{-1}{u^2} = \frac{1}{u} \end{cases} \\
 &= \iint_{v \leq x} f_{(V,W)}\left(u, \frac{x}{u}\right) \cdot \left|\frac{1}{u}\right| du dv \\
 &= \int_{-\infty}^x dv \int_{-\infty}^{+\infty} f_{(V,W)}\left(u, \frac{x}{u}\right) \cdot \left|\frac{1}{u}\right| du
 \end{aligned}$$

$$\boxed{f_Z(x) = \frac{\delta F_Z(x)}{\delta x} = \int_{-\infty}^{+\infty} f_{(V,W)}\left(u, \frac{x}{u}\right) \cdot \left|\frac{1}{u}\right| du}$$

ou si indépendant

$$\boxed{f_Z(x) = \frac{\delta F_Z(x)}{\delta x} = \int_{-\infty}^{+\infty} f_V(u) \cdot f_W\left(\frac{x}{u}\right) \cdot \left|\frac{1}{u}\right| du}$$

2.2.5 Distribution de la somme des espérances $E(V + W)$

$$\boxed{E(V + W) = E(V) + E(W)}$$

Demonstration

$$\begin{aligned}
 E(Z) &= \int_{-\infty}^{\infty} x f_{(V+W)}(x) dx \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{(V,W)}(u, x-u) du dv && \text{2.2.3 page 13} \\
 \text{On remplace par } &\left\{ \begin{array}{l} u = \xi \\ x = \xi - \eta \\ J = \begin{pmatrix} \frac{\delta u}{\delta \xi} & \frac{\delta u}{\delta \eta} \\ \frac{\delta x}{\delta \xi} & \frac{\delta x}{\delta \eta} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} = 1 \cdot -1 - 0 \cdot 1 = -1 \end{array} \right. \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\xi + \eta) f_{(V,W)}(\xi, \eta) |-1| d\xi d\eta \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \xi f_{(V,W)}(\xi, \eta) d\xi d\eta + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \eta f_{(V,W)}(\xi, \eta) d\xi d\eta \\
 &= \int_{-\infty}^{\infty} \xi f_V(\xi) d\xi + \int_{-\infty}^{\infty} \eta f_W(\eta) d\eta \\
 &= E(V) + E(W)
 \end{aligned}$$

2.2.6 Distribution du produit des espérances $E(V.W)$

$$E(V.W) = E(V).E(W)$$

Demonstration

Uniquement pour des variables indépendantes

$$\begin{aligned}
 E(Z) &= \int_{-\infty}^{\infty} x f_{(V,W)}(x) dx \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_V(u) f_W\left(\frac{x}{u}\right) \left|\frac{1}{u}\right| du dv && \text{2.2.4 page 14} \\
 \text{On remplace par } &\begin{cases} u = \xi \\ x = \xi\eta \\ J = \begin{pmatrix} \frac{\partial u}{\partial \xi} & \frac{\partial u}{\partial \eta} \\ \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \eta & \xi \end{pmatrix} = 1.\xi - 0.\eta = \xi \end{cases} \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\xi\eta) f_{(V,W)}(\xi, \eta) \left|\frac{1}{u}\right| |\xi| d\xi d\eta \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\xi\eta) f_{(V,W)}(\xi, \eta) \left|\frac{1}{u}\right| |u| d\xi d\eta \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\xi\eta) f_{(V,W)}(\xi, \eta) d\xi d\eta \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \xi f_V(\xi) f_W(\eta) d\xi d\eta \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \eta f_{(V,W)}(\xi, \eta) d\xi d\eta \\
 &= \int_{-\infty}^{\infty} \xi f_V(\xi) d\xi \int_{-\infty}^{\infty} \eta f_W(\eta) d\eta \\
 &= E(V).E(W)
 \end{aligned}$$

2.2.7 Variance d'une variable aléatoire $D^2(aV + b)$

$$\boxed{D^2(aV + b) = a^2 D^2(V)}$$

Démonstration

D'après la propriété 2.2.1 page 11 : $\sigma_V^2 = D^2(V) = \int_{-\infty}^{\infty} (x - \mu)^2 f_V(x) dx = E((x - \mu)^2)$

$$\begin{aligned} D^2(W) &= D^2(aV + b) \\ &= E((W - (a\mu + b))^2) \\ &= E((aV + b - a\mu - b)^2) \\ &= E(a^2(V - \mu)^2) \\ &= a^2 D^2(V) \end{aligned}$$

2.2.8 Variance de la somme de deux variables aléatoires $D^2(V + W)$

$$\boxed{D^2(V + W) = D^2(V) + D^2(W) + 2\mu_{11}}$$

Démonstration

D'après la propriété 2.2.1 page 11 : $\sigma_V^2 = D^2(V) = \int_{-\infty}^{\infty} (x - \mu)^2 f_V(x) dx = E((x - \mu)^2)$

$$\begin{aligned} D^2(Z) &= D^2(V + W) \\ &= E((Z - (\mu_V + \mu_W))^2) \\ &= E(((V + W) - (\mu_V + \mu_W))^2) \\ &= E(((V - \mu_V) + (W - \mu_W))^2) \\ &= E((V - \mu_V)^2) + E((W - \mu_W)^2) + 2E((V - \mu_V)(W - \mu_W)) \\ &= D^2(V) + D^2(W) + 2\mu_{11} \end{aligned}$$

2.3 Variables aléatoires particulières

2.3.1 Schéma de relation entre les variables aléatoires

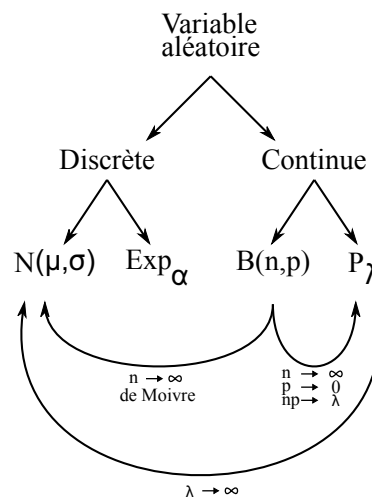


Schéma de relation entre les différentes variables aléatoires

2.3.2 Variable binomiale $\mathcal{B}(n, p)$

2.3.3 Variable de Poisson \mathcal{P}_λ

Une variable de poisson est uniquement une binomiale dont la répétition n est "grand" ($n \geq 30$) et que la probabilité p est "petit" ($p \leq 0,1$).

$$\begin{aligned}
 P(V = k) &= \binom{n}{k} p^k (1-p)^{n-k} \\
 &= \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
 &= \frac{1}{k!} n(n-1)(n-2) \dots (n-k+1) p^k (1-p)^{n-k} \\
 &= \frac{1}{k!} n^k \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) p^k (1-p)^{n-k} \\
 &= \frac{1}{k!} n^k \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) \frac{p^k (1-p)^n}{(1-p)^k} \\
 &= \frac{1}{k!} n^k \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) \frac{p^k \left(1 - \frac{np}{n}\right)^n}{(1-p)^k}
 \end{aligned}$$

Sous une autre forme :

$$\begin{aligned}
 P(V = k) &= \frac{n^k p^k}{k!} \left(1 - \frac{np}{n}\right)^n \frac{\left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right)}{(1-p)^k} \\
 &= \frac{(np)^k}{k!} e^{-\lambda} \frac{\left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right)}{(1-p)^k}
 \end{aligned}$$

Et donc quand $n \rightarrow \infty$:

$$P(V = k) \xrightarrow{n \rightarrow \infty} \frac{\lambda^k}{k!} e^{-\lambda} \frac{(1-1)(1-1) \dots (1-1)}{(1-p)^k}$$

$$\boxed{P(V = k) \xrightarrow{n \rightarrow \infty} \frac{\lambda^k}{k!} e^{-\lambda} = P_\lambda}$$

Calcul de la moyenne $E(P_\lambda)$

$$\begin{aligned}
 \mu &= \sum_{k=0}^{\infty} x_i p_i \\
 &= \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} \\
 &= \sum_{k=0}^{\infty} \frac{\lambda^{k-1} \lambda}{(k-1)!} e^{-\lambda} \\
 &= \lambda e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\
 &= \lambda e^{-\lambda} e^\lambda \\
 &= \lambda
 \end{aligned}$$

Calcul de la variance $D^2(P_\lambda)$

$$\begin{aligned}\sigma^2 &= \sum_{k=0}^{\infty} x_i^2 p_i - \mu^2 \\ &= \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} - \mu^2\end{aligned}$$

On remplace k^2 par $k(k-1) + k$

$$\begin{aligned}&= \sum_{k=0}^{\infty} k(k-1) \frac{\lambda^k}{k!} e^{-\lambda} + \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} - \mu^2 \\ &= \sum_{k=0}^{\infty} k(k-1) \frac{\lambda^k}{k!} e^{-\lambda} + \mu - \mu^2 \\ &= \sum_{k=0}^{\infty} k(k-1) \frac{\lambda^{k-2} \lambda^2}{k(k-1)(k-2)!} e^{-\lambda} + \mu - \mu^2 \\ &= \lambda^2 e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} + \mu - \mu^2 \\ &= \lambda^2 e^{-\lambda} e^{\lambda} + \mu - \mu^2 \\ &= \lambda^2 + \lambda - \lambda^2 \\ &= \lambda\end{aligned}$$

Calcul de la fonction génératrice $\psi(t)$

$$\begin{aligned}(t) &= \sum_{k=0}^{\infty} e^{tk} e^{-\lambda} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k}{k!} \\ &= e^{-\lambda} e^{\lambda e^t} \\ &= e^{\lambda(e^t - 1)}\end{aligned}$$

Calcul de la stabilité

$$\begin{aligned}\psi_{1+2}(t) &= \psi_1(t) \psi_2(t) \\ &= e^{\lambda_1(e^t - 1)} e^{\lambda_2(e^t - 1)} \\ &= e^{\lambda_1(e^t - 1)(\lambda_1 + \lambda_2)}\end{aligned}$$

La variable est bien stable.

2.3.4 Variable exponentielle négative**2.3.5 Variable Normale $\mathcal{N}(\mu, \sigma)$** **Propriété d'opposé**

Soit \bar{X}_1 et \bar{X}_2 deux variables normales. Si \bar{X}_2 est l'opposé de \bar{X}_1 ($\bar{X}_2 = -\bar{X}_1$) alors

$$\boxed{\bar{X}_1 \sim N(\mu_1, \sigma) \text{ et } \bar{X}_2 \sim N(-\mu_2, \sigma)}$$

Propriété d'addition

Soit $N(\mu_1, \sigma_1)$ et $N(\mu_2, \sigma_2)$ deux variables normales indépendantes et $\psi_1(t)$ et $\psi_2(t)$ leurs fonctions génératrices.

$$\begin{aligned} \psi(t) &= \psi_1(t) \cdot \psi_2(t) \\ &= e^{\mu_1 + t + \sigma_1^2 t^2 / 2} \cdot e^{\mu_2 + t + \sigma_2^2 t^2 / 2} \\ &= e^{(\mu_1 + \mu_2)t + \frac{1}{2}(\sigma_1^2 + \sigma_2^2)t^2} \end{aligned}$$

On obtient bien une fonction génératrice d'une normale de paramètres $(\mu_1 + \mu_2)$ et $\sqrt{\sigma_1^2 + \sigma_2^2}$.

$$\boxed{N(\mu_1, \sigma_1) + N(\mu_2, \sigma_2) = N\left(\mu_1 + \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2}\right)}$$

2.3.6 Variable Khi²**2.3.7 Variable Student t_n** **2.3.8 Variable Snedecor $\mathcal{F}_{(m,n)}$**

2.4 Théorèmes fondamentaux

2.4.1 Inégalité de Bienaymé-Tchebycheff

La proportion d'individus s'écartant de la moyenne d'une distribution de plus k fois l'écart-type (σ) ne dépasse jamais $\frac{1}{k^2}$:

$$\boxed{\frac{1}{k^2} \geq P(|V - \mu| \geq k\sigma)}$$

Démonstration

Cas continu

$$\begin{aligned} \sigma^2 &= \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \\ &= \underbrace{\int_{-\infty}^{\mu - k\sigma} (x - \mu)^2 f(x) dx}_{\substack{x \leq \mu - k\sigma \\ x - \mu \leq -k\sigma \\ (x - \mu)^2 \geq (k\sigma)^2}} + \underbrace{\int_{\mu - k\sigma}^{\mu + k\sigma} (x - \mu)^2 f(x) dx}_{\geq 0} + \underbrace{\int_{\mu + k\sigma}^{\infty} (x - \mu)^2 f(x) dx}_{\substack{x \geq \mu + k\sigma \\ x - \mu \geq k\sigma \\ (x - \mu)^2 \geq (k\sigma)^2}} \\ \sigma^2 &\geq \int_{-\infty}^{\mu - k\sigma} (k\sigma)^2 f(x) dx + \int_{\mu + k\sigma}^{\infty} (k\sigma)^2 f(x) dx \\ &\geq (k\sigma)^2 \int_{-\infty}^{\mu - k\sigma} f(x) dx + (k\sigma)^2 \int_{\mu + k\sigma}^{\infty} f(x) dx \\ \frac{1}{k^2} &\geq P(V - \mu \leq -k\sigma) + P(V - \mu \geq k\sigma) \\ &\geq P(|V - \mu| \geq k\sigma) + P(|V - \mu| \geq k\sigma) \end{aligned}$$

$$\boxed{\frac{1}{k^2} \geq P(|V - \mu| \geq k\sigma)}$$

Cas discret

$$\begin{aligned} \sigma^2 &= \sum_i p_i (x_i - \mu)^2 \\ &= \sum_{i; x_i \leq \mu - k\sigma} p(x_i - \mu)^2 + \sum_{i; \mu - k\sigma < x_i < \mu + k\sigma} p(x_i - \mu)^2 + \sum_{i; \mu + k\sigma \leq x_i} p(x_i - \mu)^2 \\ \sigma^2 &\geq \sum_{i; x_i \leq \mu - k\sigma} p(x_i - \mu)^2 + \sum_{i; \mu + k\sigma \leq x_i} p(x_i - \mu)^2 \\ &\geq k^2 \sigma^2 \sum_{i; x_i \leq \mu - k\sigma} p_i + \sum_{i; \mu + k\sigma \leq x_i} p_i \\ \frac{1}{k^2} &\geq \sum_{i; x_i \leq \mu - k\sigma} p_i + \sum_{i; \mu + k\sigma \leq x_i} p_i \\ &\geq P(V \leq \mu - k\sigma) + P(V \geq \mu + k\sigma) \\ &\geq P(V - \mu \leq -k\sigma) + P(V - \mu \geq k\sigma) \\ &\geq P(|V - \mu| \geq k\sigma) + P(|V - \mu| \geq k\sigma) \end{aligned}$$

$$\boxed{\frac{1}{k^2} \geq P(|V - \mu| \geq k\sigma)}$$

2.4.2 Théorème de Bernoulli ou loi des grands nombres

Lors de n répétitions d'une expérience aléatoire, la fréquence relative $\frac{F}{n}$ d'un évènement tend vers sa probabilité p d'exister lorsque $n \rightarrow \infty$

$$\frac{p(1-p)}{n\epsilon^2} \geq P\left(\left|\frac{F}{n} - p\right| \geq \epsilon\right) \xrightarrow{n \rightarrow \infty} 0$$

Démonstration

On part avec le théorème de Bienaymé-Tchebycheff

$$\frac{1}{k^2} \geq P(|V - \mu| \geq k\sigma)$$

Et on considère une binomiale $V = B(n, p)$

$$\begin{aligned} \frac{1}{k^2} &\geq P(|B(n, p) - np| \geq k\sqrt{np(1-p)}) \\ &\geq P\left(\frac{|B(n, p) - np|}{n} \geq \frac{k\sqrt{np(1-p)}}{n}\right) \\ &\geq P\left(\left|\frac{B(n, p)}{n} - p\right| \geq k\sqrt{\frac{np(1-p)}{n^2}}\right) \\ &\geq P\left(\left|\frac{B(n, p)}{n} - p\right| \geq k\sqrt{\frac{p(1-p)}{n}}\right) \end{aligned}$$

On pose $k\sqrt{\frac{p(1-p)}{n}} = \epsilon$ et $B(n, p) = F$

$$\begin{cases} k\sqrt{\frac{p(1-p)}{n}} &= \epsilon \\ k &= \epsilon \frac{\sqrt{n}}{\sqrt{p(1-p)}} \\ k^2 &= \frac{p(1-p)}{\epsilon^2 n} \\ \frac{1}{k^2} &= \frac{p(1-p)}{\epsilon^2 n} \end{cases}$$

$$\frac{p(1-p)}{n\epsilon^2} \geq P\left(\left|\frac{F}{n} - p\right| \geq \epsilon\right) \xrightarrow{n \rightarrow \infty} 0$$

2.4.3 Théorème Central-Limite

Ce théorème stipule que si V est une somme de n variables aléatoires (quelconques) indépendantes

$$V = X_1, X_2, X_3, \dots, X_n$$

alors sa variable réduite $\frac{V - \mu}{\sigma}$ tend vers une gaussienne $N(0, 1)$ lorsque $n \rightarrow \infty$.

$$\boxed{\frac{V - E(V)}{D(V)} = \frac{V - \mu}{\sigma} \xrightarrow{n \rightarrow \infty} N(0, 1)}$$

Démonstration

On pose d'abord les calculs de l'espérance et de l'écart-type de la somme de variable aléatoire V , pour un de ses éléments V_i , pour la somme de variable centrée W et pour un de ses éléments W_i :

$$\left\{ \begin{array}{ll} V &= \sum_{i=1}^n V_i & (1) \\ E(V_i) &= \hat{\mu} & (2) \\ D(V_i) &= \hat{\sigma} & (3) \\ E(V) &= \sum_{i=1}^n E(V_i) & \\ &= n\hat{\mu} & (4) \\ D(V) &= \sqrt{D^2(V)} & \\ &= \sqrt{\sum_{i=1}^n D^2(V_i)} & \\ &= \sqrt{n\hat{\sigma}^2} & \\ &= \sqrt{n} \hat{\sigma} & (5) \end{array} \right\} \left\{ \begin{array}{ll} W_i &= V_i - \hat{\mu} & (6) \\ W &= \sum_{i=1}^n W_i & \\ &= \sum_{i=1}^n (V_i - \hat{\mu}) & \\ &= \sum_{i=1}^n (V_i) - n\hat{\mu} & \\ &= V - E(V) & (7) \\ E(W_i) &= E(V_i - \hat{\mu}) & \\ &= E(V_i) - \hat{\mu} & \\ &= \hat{\mu} - \hat{\mu} = 0 & (8) \\ D^2(W_i) &= D^2(V_i - \hat{\mu}) & \\ &= D^2(V_i) & \\ &= \hat{\sigma}^2 & (9) \end{array} \right.$$

On commence avec la fonction génératrice des moments $\psi_V(t) = E(e^{tV})$

$$\begin{aligned} \psi_Z(t) &= E(e^{tZ}) \\ &= E\left(e^{\left(t \frac{V - \mu}{\sigma}\right)}\right) \\ &= E\left(e^{\left(\frac{t}{\sigma}(V - \mu)\right)}\right) \end{aligned}$$

$$\begin{aligned}
\text{On remplace par } \left\{ \begin{aligned}
V - \mu &= \sum_{i=1}^n V_i - \mu && \text{car propriété (1)} \\
&= \sum_{i=1}^n V_i - E(V) && \text{car propriété (4)} \\
&= \sum_{i=1}^n V_i - E\left(\sum_{i=1}^n V_i\right) && \text{car propriété (1)} \\
&= \sum_{i=1}^n V_i - \sum_{i=1}^n E(V_i) \\
&= \sum_{i=1}^n V_i - \sum_{i=1}^n \hat{\mu} && \text{car propriété (2)} \\
&= \sum_{i=1}^n V_i - \hat{\mu}
\end{aligned} \right. \\
\\
= E \left(e^{\left(\frac{t}{\sigma} \sum_{i=1}^n (V_i - \hat{\mu}) \right)} \right) \\
= E \left(e^{\left(\frac{t}{\sigma} \sum_{i=1}^n (W_i) \right)} \right) \quad \text{car propriété (6)}
\end{aligned}$$

$$\text{Propriété : } Z = V + W (\text{indépendant}) \left\{ \begin{aligned}
\psi_Z &= E(e^{tZ}) \\
&= E(e^{t(V+W)}) \\
&= E(e^{tV} e^{tW}) \\
&= E(e^{tV}) E(e^{tW}) \\
&= \psi_V \psi_W
\end{aligned} \right.$$

$$\begin{aligned}
&= \psi_{W_1} \left(\frac{t}{\sigma} \right) \psi_{W_2} \left(\frac{t}{\sigma} \right) \dots \psi_{W_n} \left(\frac{t}{\sigma} \right) \\
&= \left(\psi_{W_i} \left(\frac{t}{\sigma} \right) \right)^n
\end{aligned}$$

$$\boxed{\psi_Z(t) = E(e^{tZ}) = \left(\psi_{W_i} \left(\frac{t}{\sigma} \right) \right)^n}$$

Maintenant nous développons $\psi_{W_i}(t)$ par le théorème de Taylor :

$$\left\{ \begin{aligned}
f(x) &= \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + R_n(x) \\
&= f(a) + \frac{f'(a)}{1!} (x-a) + \frac{f^{(2)}(a)}{2!} (x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!} (x-a)^n + R_n(x)
\end{aligned} \right.$$

$$\boxed{\begin{aligned}
D^2(W_i) &= E(W_i^2) - E^2(W_i) \\
E(W_i^2) &= D^2(W_i) + E^2(W_i) \\
&= \hat{\sigma}^2 + 0^2
\end{aligned}}$$

$$\left\{ \begin{aligned}
\psi_{W_i}(t) &= 1 + E(W_i)t + E(W_i^2)\frac{t^2}{2} + R(t^3) \\
&= 1 + 0t + \hat{\sigma}^2\frac{t^2}{2} + R(t^3) \quad \text{car propriété (8) et (9)}
\end{aligned} \right.$$

$$\boxed{Z = \frac{W}{\sqrt{n\hat{\sigma}}} = \frac{V - \mu}{\sigma} \xrightarrow{n \rightarrow \infty} N(0, 1)}$$

$$\begin{aligned} \lim_{n \rightarrow \infty} \psi_Z(t) &= \lim_{n \rightarrow \infty} \left(\psi_{W_i} \left(\frac{t}{\sqrt{n\hat{\sigma}}} \right)^n \right) \\ &= \lim_{n \rightarrow \infty} \left(1 + \frac{t^2}{2n} + R(t^3) \right)^n \\ &= \lim_{n \rightarrow \infty} \left(1 + \frac{t^2}{2n} \right)^n \\ &= e^{\frac{t^2}{2}} \end{aligned}$$

On a donc une allure d'une fonction génératrice d'une $N(0, 1)$

$$\boxed{\psi_Z(t) \xrightarrow{n \rightarrow \infty} N(0, 1)}$$

2.4.4 Théorème de De Moivre

C'est un cas particulier du théorème Central-Limite puisqu'une binomiale est bien une somme de variables quelconques de mêmes distributions (à savoir, des variables indicatrices). La variable binomiale est asymptotiquement normale lorsque $n \rightarrow \infty$.

$$\boxed{\frac{B(n, p) - np}{\sqrt{np(1-p)}} \xrightarrow[n \rightarrow \infty]{} N(0, 1)}$$

Démonstration

On part donc avec le théorème de Central-Limite :

$$\frac{V - E(V)}{D(V)} = \frac{V - \mu}{\sigma} \xrightarrow[n \rightarrow \infty]{} N(0, 1)$$

Et on considère une binomiale $V = B(n, p)$.

$$\boxed{\frac{B(n, p) - np}{\sqrt{np(1-p)}} \xrightarrow[n \rightarrow \infty]{} N(0, 1)}$$

3 L'inférence statistique

3.1 Les valeurs

Échantillon de la normale $\frac{n_1 S_1^2}{\sigma^2}$

Variance de l'échantillon S_1^2

Variance théorique de l'ensemble de la population σ^2

3.2 Distributions échantillonnées

3.2.1 Distribution échantillonnées de la moyenne \bar{X}

Supposons que la distribution de la population ait une moyenne μ et un écart-type σ . Il en est donc de même pour les distributions échantillonnées de X_1, X_2, \dots, X_n et il résulte alors des propriétés de l'espérance mathématique que :

$$\begin{aligned} E(\bar{X}) &= E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) \\ &= \frac{1}{n} \sum_{i=1}^n E(X_i) \\ &= \frac{1}{n} \sum_{i=1}^n \mu \\ &= \mu \end{aligned}$$

et il résulte des propriétés de la variance que :

$$D^2(\bar{X}) = D^2\left(\frac{1}{n} \sum_{i=1}^n X_i\right)$$

$$D^2(aV + b) = a^2 D^2(V) \quad 2.2.7 \text{ page 17}$$

$$\begin{aligned} &= \frac{1}{n^2} \sum_{i=1}^n D^2(X_i) \\ &= \frac{1}{n^2} \sum_{i=1}^n \sigma^2 \\ &= \frac{\sigma^2}{n} \end{aligned}$$

3.2.2 Distribution échantillonnées de la variance S^2

$$\begin{aligned} E(S^2) &= E\left(\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2\right) \\ &= \frac{1}{n} \sum_{i=1}^n E(X_i^2) + 2E\left(\frac{1}{n} \sum_{i=1}^n X_i \bar{X}\right) + \frac{1}{n} \sum_{i=1}^n E(\bar{X}^2) \end{aligned}$$

3.2.3 Distribution échantillonnées d'une fonction de \bar{X} et s^2

Il résulte de ce qui précède que lorsque la population a une distribution normale, la variable

$$\sqrt{n-1} \frac{\bar{X} - \mu}{S} \sim t_{(n-1)}$$

Démonstration

La définition d'une chi carré $\chi^2_{(n-1)}$ est une somme de normale au carré $\sum_{i=1}^{n-1} (N_i(0, 1))^2$.

Or on a vu que $\chi^2_{(n-1)} = \frac{nS^2}{\sigma^2}$, donc

$$\begin{aligned} \chi^2_{(n-1)} &= \sum_{i=1}^{n-1} (N_i(0, 1))^2 \\ \frac{nS^2}{\sigma^2} &= \sum_{i=1}^{n-1} \left(\frac{N_i(0, \sigma)}{\sigma} \right)^2 \\ \frac{nS^2}{\sigma^2} &= \sum_{i=1}^{n-1} \frac{(N_i(0, \sigma))^2}{\sigma^2} \\ nS^2 &= \sum_{i=1}^{n-1} (N_i(0, \sigma))^2 \\ nS^2 \frac{1}{n-1} &= \frac{1}{n-1} \sum_{i=1}^{n-1} (N_i(0, \sigma))^2 \\ \left(\frac{nS^2}{n-1} \right)^{\frac{-1}{2}} &= \left(\frac{1}{n-1} \sum_{i=1}^{n-1} (N_i(0, \sigma))^2 \right)^{\frac{-1}{2}} \\ \frac{1}{\sqrt{\frac{nS^2}{n-1}}} &= \frac{1}{\sqrt{\frac{1}{n-1} \sum_{i=1}^{n-1} (N_i(0, \sigma))^2}} \\ \frac{N(0, \sigma)}{S \sqrt{\frac{n}{n-1}}} &= \frac{N(0, \sigma)}{\sqrt{\frac{1}{n-1} \sum_{i=1}^{n-1} (N_i(0, \sigma))^2}} \\ \frac{N(0, \sigma)}{S \sqrt{\frac{n}{n-1}}} &= t_{(n-1)} \end{aligned}$$

On obtient bien une student $t_{(n-1)}$ à $n-1$ degrés de liberté (2.3.7 page 21). On va donc remplacer la normale $N(0, \sigma)$ dans l'équation par :

$$\begin{cases} \bar{X} & \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \\ \frac{\bar{X} - \mu}{\frac{1}{\sqrt{n}}} & \sim N(0, \sigma) \\ \sqrt{n}(\bar{X} - \mu) & \sim N(0, \sigma) \end{cases}$$

$$t_{(n-1)} = \frac{N(0, \sigma)}{S \sqrt{\frac{n}{n-1}}} = \frac{\sqrt{n}(\bar{X} - \mu)}{S \sqrt{\frac{n}{n-1}}} = \frac{\bar{X} - \mu}{S \frac{1}{\sqrt{n-1}}} = \sqrt{n-1} \frac{\bar{X} - \mu}{S}$$

$$\boxed{t_{(n-1)} = \sqrt{n-1} \frac{\bar{X} - \mu}{S}}$$

3.3 Test d'hypothèse

Erreur de première espèce est de rejeter une hypothèse alors qu'elle est vraie. Le risque α de commettre cette erreur est ϵ ;

Erreur de seconde espèce est d'accepter une hypothèse alors qu'elle est fausse. Si on accepte une moyenne μ_0 alors qu'elle vaut en réalité μ_1 , alors le risque de commettre cette erreur est

$$\begin{aligned}\beta &= P\left(\mu_0 - u_{\beta/2} \frac{\sigma}{\sqrt{n}} \leq \bar{X} \leq \mu_0 + u_{\beta/2} \frac{\sigma}{\sqrt{n}}\right) \\ &\left\{ \begin{array}{l} \text{où } \bar{X} \sim N\left(\mu_1, \frac{\sigma}{\sqrt{n}}\right) \Rightarrow \frac{\bar{X} - \mu_1}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1) \end{array} \right. \\ &= P\left(\mu_0 - u_{\beta/2} \frac{\sigma}{\sqrt{n}} \leq N(0, 1) \frac{\sigma}{\sqrt{n}} + \mu_1 \leq \mu_0 + u_{\beta/2} \frac{\sigma}{\sqrt{n}}\right) \\ &= P\left(\frac{\mu_0 - \mu_1}{\sigma} \sqrt{n} - u_{\beta/2} \leq N(0, 1) \leq \frac{\mu_0 - \mu_1}{\sigma} \sqrt{n} + u_{\beta/2}\right)\end{aligned}$$

Puissance d'un test est la valeur $1 - \beta$, c'est à dire la probabilité de rejeter l'hypothèse alors qu'elle est fausse.

3.4 Intervalles de confiance et tests d'hypothèses

3.4.1 Comparaison des moyennes de deux populations normales de deux écart-type σ connu

On commence avec deux moyennes des deux populations :

$$\bar{X}_1 \sim N\left(\mu_1, \frac{\sigma_1}{\sqrt{n_1}}\right)$$

$$\bar{X}_2 \sim N\left(\mu_2, \frac{\sigma_2}{\sqrt{n_2}}\right)$$

Maintenant nous devons faire une soustraction de deux normales $\bar{X}_1 - \bar{X}_2$, pour faire cela nous prenons l'opposé d'une normale et nous les additionnons $\bar{X}_1 + (-\bar{X}_2)$ grace à la propriété d'opposé d'une normale 2.3.5 page 21 et la propriété d'addition 2.3.5 page 21.

$$\begin{aligned}\bar{X}_1 + (-\bar{X}_2) &\sim N\left(\mu_1 - \mu_2, \sqrt{\left(\frac{\sigma_1}{\sqrt{n_1}}\right)^2 + \left(\frac{\sigma_2}{\sqrt{n_2}}\right)^2}\right) \\ &\sim N\left(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)\end{aligned}$$

Ces deux variables étant indépendantes, on en déduit que

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1)$$

A tout $\epsilon > 0$, on peut donc associé un nombre tel que

$$P(-u_{\epsilon/2} \leq T \leq u_{\epsilon/2}) = 1 - \epsilon$$

d'où on peut donc déduire un intervalle de confiance pour $\mu_1 - \mu_2$. D'autre part on rejettera l'hypothèse $\mu_1 = \mu_2$ si

$$\frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > u_{\epsilon/2}$$

3.4.2 Comparaison des moyennes de deux populations normales de même écart-type σ inconnu

On commence avec deux moyennes des deux populations :

$$\bar{X}_1 \sim N\left(\mu_1, \frac{\sigma}{\sqrt{n_1}}\right)$$

$$\bar{X}_2 \sim N\left(\mu_2, \frac{\sigma}{\sqrt{n_2}}\right)$$

Maintenant nous devons faire une soustraction de deux normales $\bar{X}_1 - \bar{X}_2$, pour faire cela nous prenons l'opposé d'une normale et nous les additionnons $\bar{X}_1 + (-\bar{X}_2)$ grâce à la propriété d'opposé d'une normale 2.3.5 page 21 et la propriété d'addition 2.3.5 page 21.

$$\begin{aligned}\bar{X}_1 + (-\bar{X}_2) &\sim N\left(\mu_1 - \mu_2, \sqrt{\left(\frac{\sigma}{\sqrt{n_1}}\right)^2 + \left(\frac{\sigma}{\sqrt{n_2}}\right)^2}\right) \\ &\sim N\left(\mu_1 - \mu_2, \sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}\right) \\ &\sim N\left(\mu_1 - \mu_2, \sigma \sqrt{\frac{n_1 + n_2}{n_1 n_2}}\right)\end{aligned}$$

et avec un $\chi^2_{(n_1-1+n_2-1)}$ où n = nombre de degré de libertés

$$\begin{aligned}\frac{n_1 s_1^2 + n_2 s_2^2}{\frac{1}{\sigma} \sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} &\sim \chi^2_{(n_1+n_2-2)} \\ \frac{1}{\sigma} \sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}} &\sim \sqrt{\chi^2_{(n_1+n_2-2)}}\end{aligned}$$

Ces deux variables étant indépendantes, on en déduit que

$$T = \underbrace{\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sigma \sqrt{\frac{n_1 + n_2}{n_1 n_2}}}}_{\sim N(0,1)} \underbrace{\frac{1}{\frac{1}{\sigma} \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}}}_{\sim \chi^2_{n_1-1+n_2-1}}$$

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{n_1 + n_2}{n_1 n_2}} \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}} \sim t_{n_1-1+n_2-1}$$

A tout $\epsilon > 0$, on peut donc associer un nombre tel que

$$P(-t_{\epsilon/2} \leq T \leq t_{\epsilon/2}) = 1 - \epsilon$$

d'où on peut donc déduire un intervalle de confiance pour $\mu_1 - \mu_2$. D'autre part on rejettera l'hypothèse $\mu_1 = \mu_2$ si

$$\frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{n_1 + n_2}{n_1 n_2}} \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}} > t_{\epsilon/2}$$

3.4.3 Comparaison des moyennes de deux populations quelconques

Si n_1 et n_2 sont suffisamment grands (au moins 20), alors $\bar{X}_1 - \bar{X}_2$ a une distribution approximativement normale :

$$\bar{X}_1 - \bar{X}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

où on remplace σ_1 et σ_2 par s_1 et s_2 lorsqu'ils sont inconnus. On rejette alors l'hypothèse $\mu_1 = \mu_2$ lorsque

$$\frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} > u_{\epsilon/2}$$

4 Autres

4.1 Tableau du formulaire

	μ	σ^2	$\psi(t)$
$\mathcal{B}(n, p)$	np	$np(1-p)$	$(pe^t + q)^n$
\mathcal{P}_λ	λ	λ	$e^{\lambda(e^t-1)}$
Exp_λ	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda-t}$
Indicatrice(p)	p	$p(1-p)$	$1 + p(e^t - 1)$
Uniforme[a, b]	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{1}{t} \frac{e^{tb} - e^{ta}}{b-a}$
$\mathcal{N}(\mu, \sigma)$	μ	σ^2	$e^{\mu t + (\sigma^2 t^2)/2}$
$\chi_{(n)}^2$	n	$2n$	$(1 - 2t)^{-n/2}$
t_n	$0 \quad n > 1$	$\frac{n}{n-2} \quad n > 2$	aucun
$\mathcal{F}_{(m,n)}$	$\frac{n}{n-2} \quad n > 2$	$\frac{2n^2(n+m-2)}{m(n-2)^2(m-4)} \quad n > 2$	aucun

Tableau dans le formulaire disponible à l'examen écrit (**en rouge à connaître**)

4.2 Densité et répartition

	Fonction de densité $f(x)$	Fonction de répartition $F(x)$
$\mathcal{B}(n, p)$	$P[B(n, p) = k]$	$\sum_{k=0}^x P[B(n, p) = k]$
\mathcal{P}_λ	$P[\mathcal{P}_\lambda = k] = \frac{\lambda^k}{k!} e^{-\lambda}$	$\sum_{k=0}^x P[\mathcal{P}_\lambda = k]$
Exp_λ	$\begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$	$\begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$
Indicatrice(p)	$V_A \Rightarrow \begin{cases} P(V_A = 1) = p \\ P(V_A = 0) = 1 - p \end{cases}$	$\begin{cases} 0 & x < 0 \\ 1 - p & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$
Uniforme[a, b]	$\begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{sinon} \end{cases}$	$\begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x < b \\ 1 & x \geq b \end{cases}$
$\mathcal{N}(\mu, \sigma)$	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(u-\mu)^2}{2\sigma^2}} du$
$\chi_{(n)}^2$	$\begin{cases} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} & x > 0 \\ 0 & x \leq 0 \end{cases}$	$\frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{u^2}{2}} du$
t_n	Densité indépendante de σ	
$\mathcal{F}_{(m,n)}$	Densité indépendante de σ	

4.3 Distributions

