Corrigés des TD de Statistiques

Série 1 de TD

Exercice1 : Une entreprise utilise une matière isolante dans l'assemblage de certains appareils de mesure de contrôle industriel. Ces composantes isolantes sont achetées d'un fournisseur américain et doivent respecter une certaine épaisseur. Lors d'un contrôle de réception, on a mesuré l'épaisseur d'un échantillon de vingt composantes :

Epaisseur en mm				
5,6	5,9	6,2	6,1	6,6
5,9	5,9	5,6	6,2	5,8
5,5	5,6	6,0	6,3	6,2
5,9	6,2	6,0	6,2	6,3

1. Calculer l'épaisseur moyenne de cet échantillon
2. Quelle est l'étendue de cette série statistique ?
3. Calculer la variance et l'écart-type de l'épaisseur des composantes isolantes.
4. Un lot est considéré comme acceptable si l'épaisseur moyenne observée dans un échantillon de 20 n'est pas inférieure à 5,8 mm, ni supérieure à 6,2 mm. Devrait-on retourner ce lot au fournisseur ?

Exercice 2 : Chaque jour, pour aller travailler, un employé a le choix entre les transports publics ou son véhicule. Un échantillon des temps de trajets avec chacune des deux méthodes est présenté ci-dessous. Les temps sont exprimés en minutes.

Transport public :	28	29	32	37	33	25	29	32	41	34
Véhicule :	29	31	33	32	34	30	31	32	35	33

1. Calculer le temps moyen du trajet pour aller travailler effectué avec chacune des deux méthodes de transport
2. Calculer l'écart-type pour les deux méthodes.

Sur la base de vos résultats aux questions 1) et 2), quelle méthode de transport préconiseriez –vous ? Expliquer

Exercice 3 : Une entreprise se spécialisant dans la vente d'articles de sport possède 72 points de vente répartis au Nord du Maroc (40 points de vente) et en Sud du Maroc (32 points de vente). Le service de comptabilité de l'entreprise a mis en main le chiffre d'affaires de chaque point de vente pour le mois de décembre.

a) Calculer le chiffre d'affaires moyen au cours du mois de décembre dans chaque province.

b) Déterminer la variance et l'écart-type du chiffre d'affaires réalisé dans chaque province.

c) Est-ce que le chiffre d'affaires présente sensiblement le même étalement dans chaque province ?

d) Déterminer le chiffre d'affaires total moyen réalisé dans les deux provinces ?

Chiffre d'affaires- Nord du Maroc							
9016	9551	10179	9070	10220	8859	9460	9549
9393	9502	9219	9825	9845	9417	9345	10037
9852	9627	9771	9897	10140	10180	9186	8724
9729	9877	9370	9890	9688	9188	9107	9130
9118	9675	9286	9388	8297	8829	9595	9553

Chiffre d'affaires- Sud du Maroc							
10024	9936	9994	10188	10652	10266	10387	9879
10310	10510	10198	9947	10303	10237	9851	9973
10554	10107	10130	10056	10015	10520	9856	9709
9851	10063	9878	9487	10644	10755	10540	9925

Exercice 4 : Le tableau suivant représente la répartition des 1049 individus d'après la réponse à la question suivante : Idéalement, vers quel âge souhaiterez –vous prendre votre retraite ?

Peut-on calculer l'âge moyen anticipé pour la retraite ? Expliquer.

	Nombre
Le plutôt possible, maintenant	6
A 40 ans ou avant	86
A 50 ans ou avant (mais plus de 40 ans)	293
A 55 ans ou avant (mais plus de 50 ans)	272
A 60 ans ou avant (mais plus de 55 ans)	247
A 65 ans ou avant (mais plus de 60 ans)	96
A 66 ans et plus	23
Jamais, le plus tard possible	26

Exercice 5 : Un entrepreneur en construction doit ériger une structure de béton armé. La résistance en compression tel que précisé dans le contrat, pour le mélange que l'on doit utiliser, ne doit pas avoir une résistance à la compression inférieur à 300 kg/cm^2 à l'age de 28 jours. Pour assurer que cette norme soit respectée par l'usine Bétonmix, fournisseur de béton on a effectué des essais de résistance en compression sue des cylindres standards. Les résultats basés sur vingt- cinq échantillons de trois cylindres chacun (observations effectuées de la même façon et dans les mêmes conditions) sont présentés dans la distribution de fréquences absolues suivante. Distribution de fréquences absolues de la résistance à la compression à l'âge de 28 jours → effectif

Classes	Nombre de cylindres
$270 \le x < 290$	2
$290 \le x < 310$	5
$310 \le x < 330$	15
$330 \le x < 350$	19
$350 \le x < 370$	20
$370 \le x < 390$	9
$390 \le x < 410$	5

a) A l'aide des formules simplifiées, calculer la résistance moyenne à la compression, la variance et l'écart-type

b) On considère que le mélange du béton est de bonne qualité si le coefficient de variation de la résistance à la compression n'est pas supérieur à 10 %. Que peut-on conclure de la qualité du béton fourni par Bétonmix ?

Interpolation linéaire ⇐ c) Pour assurer que le béton n'ait pas, dans la majorité des cas, une résistance inférieure à 300 kg/cm^2. On a établi que l'écart entre cette norme et la résistance moyenne à la compression des essais ne doit pas être inférieur à 40 kg/cm^2. Peut-on conclure que le contracteur va être en mesure de respecter la norme spécifiée sur le contrat ?

d) Dresser un tableau indiquant les fréquences cumulées croissantes et calculer la valeur médiane de la résistance à la compression de ces essais. Que représente cette valeur ?

Exercice 6 :

Le tableau suivant fournit le pourcentage de femmes travaillant dans chaque société (x) et le pourcentage de poste à responsabilité occupés par des femmes dans chaque société (y).

Sociétés	X	Y
Coca Cola	72	61
General Motors	47	16
McDonald's	51	32
Microsoft	57	46
Dell	55	36

1. Représenter le nuage de points associés à ces données.
2. Quelle relation entre x et y le nuage de points indique-t-il ?
3. Estimer l'équation de la régression obtenue avec ces données.
4. Prévoir le pourcentage de postes à responsabilité occupés par des femmes dans une société comptant 60% de femmes parmi ses employés.
5. Utiliser l'équation de sa régression pour prévoir le pourcentage de postes à responsabilité détenus par des femmes. Comparer cette valeur aux 36 % observée chez Dell, une société qui compte 55% de femmes parmi ses employés.

05-10-2015

Série 1 de TD

Exercice 3

	Nord	Sud
a) Moyenne	9489,6	10148,2813
Var	183 043,39	89578,1396
Ecart type	427,83	299,2960
Coeff var	4,51 %	2,95 %

c) En terme de chiffre d'affaire, le sud est plus homogène que le nord.

d) le chiffre d'affaire total moyen :

$$\bar{X}_T = \frac{40\,\bar{X}_N + 32\,\bar{X}_S}{72}$$

A.N $\bar{X}_T = \frac{40 \times \bar{X}_N + 32 \times \bar{X}_S}{72} = 9782,35$

Exercice 4

Non, la série est une variable qualitative

Un autre indicateur peut être appliqué qui est le mode : 293 est le nombre le plus grand donc on prend l'intervalle entre 40 et 50

Exercice 5

a) la résistance à la compression ⇒ variable quantitative continue

Classes	nbre cylindres	Centre (c_i)	$c_i * n_i$	effectifs cumulés
$270 \leq x < 290$	2	280	560	2
$290 \leq x < 310$	5	300	1500	7
$310 \leq x < 330$	15	320	4800	22
$330 \leq x < 350$	19	340	6460	41
$350 \leq x < 370$	20	360	7200	61
$370 \leq x < 390$	9	380	3420	70
$390 \leq x < 410$	5	400	2000	75

la résistance moyenne à la compression :

$$\sum_{i} \frac{c_i * n_i}{N} = \frac{560 + 1500 + 4800 + 6460 + 7200 + 3400 + 2000}{75}$$

$= 345,87$

La variance $V = \frac{1}{N} \sum_i (c_i - \bar{x})^2 n_i$

$$\text{Coeff var} = \frac{\text{écart type}}{\text{moyenne}}$$

var	786,92
ecart type	28,05
coeff var	8,11 %

b) Bonne qualité

c)

x_i	n_i
330	22
Me	37,5 (50%)
350	41

l'interpolation linéaire impliqueé :

$$\frac{Me - 330}{37{,}5 - 22} = \frac{350 - 300}{41 - 22}$$

$\Rightarrow$ Me = 370,8

Exercice 6

1) Coeff correlat° = 0,95 Relat° ascendante

$Y = aX + b$

$$a = \frac{COV(x,y)}{V(x)}$$

$b = \bar{y} - a\bar{x}$

4) 44,18 %

5) 35,845

ENSIAS
$2^{ème}$ année: 2015/2016

Statistique
Mr S. ACHCHAB

SÉRIE:2

Echantillonnage et Estimation ponctuelle

Exercice 1

Un commerçant propose à sa clientèle six articles électroménagers. Considérons la population mère constituée par ces six articles codés ω_i ($i \in 1, ..., 6$). Soit X la variable qui représente "le nombre d'unités en stock de chaque article au moment de l'inventaire".

Ω	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6
X	0	1	2	3	0	1

1. Détreminer la loi de probabilité de la variable aléatoire X. Calculer la moyenne $\mu = E(X)$ et la variance $\sigma^2 = Var(X)$, ainsi que σ^4 et le moment centré $\mu_4 = E((X-\mu)^4)$.

2. Dans cette population d'effectif $N = 6$, on tire avec remise des échantillons de taille $n = 2$. X_1 est le nombre d'unités en stock pour le premier article tiré et X_2 est le nombre d'unités en stock pour le second. On pose

$$\overline{X} = \frac{1}{2}(X_1 + X_2).$$

Détreminer les valeurs prises par $\overline{X}$ sur tous les échantillons possibles, en déduire sa loi, calculer l'espérance empirique $E(\overline{X})$ et la variance empirique $Var(\overline{X})$. Vérifier les résultats théoriques du cours.

3. On considère la statistique S^2 définie par

$$S^2 = \frac{1}{2}\sum_{i=1}^{2}(X_i - \overline{X})^2 = \frac{1}{2}(X_1^2 + X_2^2) - \overline{X}^2$$

Déterminer les valeurs prises par S^2 sur tous les échantillons de taille $n = 2$; en déduire sa loi, son espérance et sa variance.

Exercice 2

Dans un pays les statistiques font ressortir que 64% des ménages possèdes une voiture de tourisme. Quelle est la probabilité que sur un échantillon au hasard de 225 ménages, la proportion de ceux qui possèdent une voiture soit:

(a) comprise entre 40% et 70%.

(b) supérieure à 60%.

(c) inférieure à 25%

Exercice 3

Un dispositif de signalisation lumineuse comporte trois lampes; celle qui en service est relayée automatiquement en cas de défaillance.
Quelle est la probabilité que l'ensemble fonctionne:

(a) plus de 5000 heures.

(b) moins de 4200 heures.

On sait que la durée de vie des lampes utilisée est une variable normale de moyenne 1500 heures et d'écart-type 150 heures.

Exercice 4

Un paquet de tabat produit par la régie des Tabac a un poids moyen de $50g$ et un écart-type de $2g$. En supposant que ce tabac soit livré par lots de mille paquets; quelle est la probabilité que la différence $A - B$ entre les poids de deux lots A et B excède $200g$.

Exercice 5

Lors d'un concours radiophonique, on note X le nombre de réponses reçues chaque jour. On suppose que X suit une loi normale de paramètres μ et σ. Durant les 10 jours, on a obtenu

$$x_1 = 200, \quad x_2 = 240, \quad x_3 = 190, \quad x_4 = 150, \quad x_5 = 220$$

$$x_6 = 180, \quad x_7 = 170, \quad x_8 = 230, \quad x_9 = 210, \quad x_{10} = 210$$

Donner une estimation ponctuelle de μ et σ^2.

Exercice 6

Soit $(X_1, X_2, \dots, X_n)$, un n-échantillon de loi de Bernouilli $B(p)$ et soit

$$S_n = \sum_{k=1}^{n} X_k, \qquad \bar{X}_n = \frac{S_n}{n}$$

1. Montrer que $\bar{X}_n$ est un estimateur sans biais de p.
2. On cherche à estimer la variance $\sigma^2 = p(1-p)$. On propose l'estimateur

$$U_n = \bar{X}_n(1 - \bar{X}_n).$$

   (a) Calculer l'espérance et la variance de S_n.

   (b) Montrer que U_n est un estimateur biasé de σ^2.

   (c) Donner un estimateur V_n sans bias de σ^2, fonction de U_n.

Exercice 7

A la veille d'une consultation électorale, on a interrogé cent électeurs constituant un échantillon au hasard. 64 d'entre eux ont déclaré avoir l'intention de voter pour le condidat Y.

Entre quelle limite, au moment du sondage, avec une probabilité de 0.95, la proportion du corps électoral favorable au condidat Y se situe-t-elle?

Exercice 8

Dans un centre de recrutement, on a mesuré la taille de 400 conscrits. Pour cet échantillon pris au hasard, la taille moyenne $\bar{x}$ est égale à 172 cm et l'écart-type estimé s est égal à 4 cm.

Construire un intervalle qui contienne avec une probabilité 0.99 la taille moyenne de l'ensemble des conscrits de ce centre de recrutement.

Exercice 9

Si l'écart-type de la durée de vie d'un modèle de lampe électronique est estimé à 100 heures, quelle doit être la taille de l'échantillon à prélever pour que l'erreur sur l'estimation de la durée de vie moyenne n'exède pas 20 heures et avec une probabilité

a) de 95%

b) de 99%

12-10-2015

Série 2

Exercice 1

1) la loi de probabilité de la var aléatoire X : (les proba respectives de s
les valeurs que prend la VA sont : 0, 1, 2, 3

avec	0	1	2	3
$P(X=x_i)$	$1/3$	$1/3$	$1/6$	$1/6$

Calcul de $E(x)$:

$$E(x) = \sum x_i P_i$$

$$E(x) = 1/3 + 1/3 + 1/2 = 7/6$$

Calcul de $V(x)$

$$V(x) = E(X^2) - E(x)^2 = E(X - E(X))^2$$

$$= \sum_i (x_i - E(x))^2 p_i = \sum x_i^2 p_i - E(x)^2$$

$$V(x) = 1/3 + 4/6 + 9/6 - (7/6)^2$$

$$= 15/6 - (7/6)^2 = \frac{15\times 6 - 49}{36}$$

$$V(x) = 41/36$$

$$E[(X-\mu)^4] = \sum_i (x_i - \mu)^4 p_i$$

2) 36 échantillons possibles (6x6 ← avec remise)

on a $\overline{X} = 1/2 (x_1 + x_2)$

$X_1 \backslash X_2$		ω_1	ω_2	ω_3	ω_4	ω_5	ω_6
		0	1	2	3	0	1
ω_1	0	0	1/2	1	3/2	0	1/2
ω_2	1	1/2	1	3/2	2	1/2	1
ω_3	2	1	3/2	2	5/2	1	3/2
ω_4	3	3/2	2	5/2	3	3/2	2
ω_5	0	0	1/2	1	3/2	0	1/2
ω_6	1	1/2	1	3/2	2	1/2	1

$\bar{X} = \frac{1}{2}(X_1 + X_2)$

$\bar{X}$	0	1/2	1	3/2	2	5/2	3
$P(\bar{X} = \bar{x}_i)$	4/36	8/36	8/36	8/36	5/36	2/36	1/36
$\bar{x}_i \times p_i$	0	4/36	8/36	12/36	10/36	5/36	3/36

$E(\bar{X}) = \sum x_i p_i$

$= 42/36 = 7/6 = E(X) = \mu$

$V(\bar{X}) = \sigma^2/2$ (A démontrer)

$V(\bar{X}) = \frac{41}{72}$

avec $V(\bar{X}) = E(\bar{X}^2) - E(\bar{X})^2$

$= \sum \bar{x}_i^2 p_i - (7/6)^2$

3) $S^2 = 1/2 \sum_{i=1}^{2} (x_i - \bar{x})^2$

$= 1/2 (x_1^2 + x_2^2) - \bar{x}^2$

Exercice 2

X : "Posséder ou non une voiture"

$X = \begin{cases} 1 & \text{si on possède une voiture} \\ 0 & \text{sinon} \end{cases}$

$X \sim B(p)$ où $p = 0,64$

$X_1, \dots, X_{225}$ $\quad F_n = \frac{1}{n} \sum_{i=1}^{n} X_i$

a) $P[0,4 < F_n < 0,7] = ?$

D'après T.C.L $\quad F_n \sim N\left(0,64, \sqrt{\frac{0,64 \times 0,36}{225}}\right)$

$F_n \sim N(0,64, 0,032)$

$P[0,4 < F_n < 0,7] = P\left[\frac{0,4 - 0,64}{0,032} < \frac{F_n - 0,64}{0,032} < \frac{0,7 - 0,64}{0,032}\right]$

$= \mathbb{P}[-7{,}5 < Z < 3{,}8]$

$= \Pi(3{,}8) - \Pi(-7{,}5)$

$= \Pi(3{,}8) - 1 + \Pi(7{,}5)$

Exercice 3

X_i "Durée de vie de la lampe n° i"

$X_i \sim N(1500; 150)$

soit $Y = X_1 + X_2 + X_3$ la v.a "Durée de vie du dispositif"

a) $\mathbb{P}[Y > 1500] =$

$Y \sim N(4500, 150\sqrt{3})$

$V(Y) = 3 \cdot 150^2$

$\sigma(Y) = \sqrt{3}\ 150$

ie $Y \sim N(4500, 259{,}80)$

$\mathbb{P}(Y > 5000) = \mathbb{P}\left[\frac{Y - 4500}{259{,}8} > \frac{500}{259{,}8}\right]$

$= \mathbb{P}(Z > 1{,}9) = 1 - \Pi(1{,}9) = 1 - 0{,}97128$

$= 0{,}02872$

Exercice 4

X " Poids d'un paquet de Tabac"

$X \sim L(50, 2)$

$X_1^A, \ldots, X_{1000}^A$

$n = 1000$

$A = \sum_{i=1}^{1000} X_i^A$

$B = \sum_{i=1}^{1000} X_i^B$

D'après le T.C.L

$\overline{X_A} = \frac{1}{n} \sum_{i=1} X_i \sim N(50, 2/\sqrt{1000})$

$\overline{X_B} = \frac{1}{n} \sum_{i=1}^{n} X_i \sim N(50\ 2/\sqrt{1000})$

$\mathbb{P}[(A - B) > 200] = \mathbb{P}[1000(\overline{X_A} - \overline{X_B}) > 2000]$

$= \mathbb{P}[\overline{X_A} - \overline{X_B} > 0{,}2]$

$\overline{X_A} - \overline{X_B} \sim N(0, 1/5\sqrt{5})$

$V(X - Y) = V(X) + V(Y) = \frac{8}{1000}$ $\sigma(\overline{X_A} - \overline{X_B}) = \frac{2\sqrt{2}}{10\sqrt{10}} = \frac{1}{5\sqrt{5}}$

19-10-20

Exercice 6 :

$S_n = \sum_{k=1}^{n} X_k$, $\bar{X}_n = \frac{S_n}{n}$ $T = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$

$S^{*2} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$

1- T_n A.S.B $\Leftarrow$ C.M-Q $\Rightarrow$ C.P

$\lim_{n\to+\infty} E(T_n) = \theta = p$

$\lim_{n\to+\infty} V(T_n) = 0$

$T_n = \bar{X}_n = \frac{\sum_{i=1}^{n} X_i}{n}$ $E(\bar{X}_n) = p$ $\bar{X}_n$ est sans biais $\Rightarrow$ A.S.B

$V(\bar{X}_n) = \sigma^2/n = \frac{p(1-p)}{n} \xrightarrow[n\to+\infty]{} 0$

2- $\sigma^2 = p(1-p)$, $U_n = \bar{X}_n(1-\bar{X}_n)$

a) S_n suit la loi binômiale $S_n \sim B(n,p)$

$E(S_n) = n.p$ $V(S_n) = np(1-p)$

b) Mq U_n est un estimateur biaisé de σ^2 :

$U_n = \bar{X}_n - \bar{X}_n^2$

$E(U_n) = E(\bar{X}_n) - E(\bar{X}_n^2)$ on a $V(\bar{X}_n) = E(\bar{X}_n^2) - (E(\bar{X}_n))^2$

$= p - V(\bar{X}_n) - (E(\bar{X}_n))^2$

$= p - \frac{p}{n}(1-p) - p^2$

$= p - p^2 - \frac{p}{n}(1-p)$

$= p(1-p) - \frac{p}{n}(1-p)$

$= \underbrace{p(1-p)}_{\sigma^2}(1-1/n) \neq \sigma^2$

c) Un estimateur sans biais de σ^2, en fonction de U_n :

$V_n = \frac{\bar{X}_n(1-\bar{X}_n)}{1-1/n} = \frac{n}{n-1}\bar{X}_n(1-\bar{X}_n) = \frac{n}{n-1}U_n$

Exercice 7 :

X "Avis d'un électeur à y"

$X \sim B(p)$

$X = \begin{cases} 1 & \text{si on est favorable à y} \\ 0 & \text{sinon} \end{cases}$

$X_1, X_2, \dots, X_{100}$ $X_i \begin{cases} 1 & \text{si on est favorable} \\ 0 & \text{sinon} \end{cases}$

$64\% = 0{,}64 = \frac{1}{100}\sum_{i=1}^{100} X_i = f_n$

(*) $\mathbb{P}[a < p < b] = 0{,}95$

$F_n = \frac{1}{n} \sum_{i=1}^{n} X_i \sim N\left(p ; \sqrt{\frac{p(1-p)}{n}}\right)$

T.C.L : $\sqrt{\frac{p(1-p)}{n}} \simeq \sqrt{\frac{f_n(1-f_n)}{n}}$ (n étant assez grand)

(*) s'écrit : $\mathbb{P}\left[\underbrace{\frac{f_n - b}{\sqrt{\frac{f_n(1-f_n)}{n}}}}_{-t_\alpha} < \frac{F_n - p}{\sqrt{\frac{f_n(1-f_n)}{n}}} < \underbrace{\frac{f_n - a}{\sqrt{\frac{f_n(1-f_n)}{n}}}}_{+t_\alpha}\right] = 0{,}95$

$a = f_n - t_\alpha \sqrt{\frac{f_n(1-f_n)}{n}}$ où t_α vérifie $\Pi(t_\alpha) = 1 - \alpha/2 = 0{,}975$

$b = f_n + t_\alpha \sqrt{\frac{f_n(1-f_n)}{n}}$

A partir de la table $N(0,1)$: $t_\alpha = 1{,}96$

Exercice 8

$n = 400$ $\bar{x} = 172$ cm $s = 4$ cm

$\mathbb{P}[a < m < b] = 0{,}99 = 1 - 0{,}01$ a, b ??

T.C.L $\bar{X}_n \sim N(m, \sigma/\sqrt{n})$

N étant suffisamment grand, $\sigma \simeq s$

$a = \bar{x}_n - t_\alpha\, s/\sqrt{n}$

$b = \bar{x}_n + t_\alpha\, s/\sqrt{n}$ où $\Pi(t_\alpha) = 1 - \alpha/2 = 0{,}995$

fct de répartit°

D'après la table $N(0.1)$ $t_\alpha \simeq 2{,}58$.

Exercice 9

a) $\sigma = 100$

$\mathbb{P}[|\bar{X}_n - m| < 20] = 1 - \alpha$.

$n > 30$, T.C.L $\bar{X}_n \sim N(m, \sigma/\sqrt{n})$

$\mathbb{P}\left[-\frac{20\sqrt{n}}{\sigma} < \frac{\bar{X}_n - m}{\sigma/\sqrt{n}} < \frac{20\sqrt{n}}{\sigma}\right] = 1 - \alpha$

$2\Pi\left(\frac{20\sqrt{n}}{\sigma}\right) - 1 = 1 - \alpha$

$\Pi\left(\frac{20\sqrt{n}}{\sigma}\right) = 1 - \alpha/2$

Pour $\alpha = 0{,}05$ (le risque est 5% !!)

$\Pi\left(\frac{20\sqrt{n}}{\sigma}\right) = 0{,}975 \Rightarrow \frac{20\sqrt{n}}{\sigma}\ 1{,}96 \Rightarrow n = \left(\frac{196}{20}\right)^2 \simeq 96{,}04$

b) : Pour $\alpha = 0{,}01$ $\Rightarrow \frac{20\sqrt{n}}{\sigma} = (2{,}58) \Rightarrow n = \left(\frac{258}{20}\right)^2 = 166{,}41$

Série 2 : Tests d'hypothèses

Cas 1 :

Dans une chaîne de production particulière, chaque boite pèse en moyenne 16 g. Le sous- ou le sur-remplissage est un problème sérieux et la chaîne de production sera fermée si l'un des deux cas se produit. Grâce aux données passées, on sait que σ est égal à 0.8 g. Un contrôleur de la qualité échantillonne 30 boites toutes les deux heures et décide de fermer ou non la chaîne de production pour ajustement.

1. Quelle est la règle de rejet pour un seuil de signification égal à 0.05 ?
2. Si la moyenne d'un échantillon $\bar{x}$ est égale à 16.32 g, quelle action recommanderiez-vous ?
3. Si la moyenne d'échantillon est égale à 15.82 g, quelle action préconiseriez-vous ?

Cas 2 :

Une opération particulière dans la chaîne de montage automobile nécessite, en moyenne, 2.2 minutes. A cause des effets du temps de montage d'une pièce à la fois sur les opérations d'assemblage précédentes et suivantes, il est important de conserver un temps moyen d'assemblage de 2.2 minutes. Un échantillon aléatoire de 45 pièces assemblées révèle un temps moyen de montage de 2.39 minutes, avec un écart type d'échantillon de 0.20 minute. Utiliser un seuil de signification égal à 0.02 et tester si l'opération requière en moyenne 2.2 minutes pour être exécutée.

Cas 3:

L'entreprise Simtech fabrique des tubes de verre pour l'entreprise Giscom. L'entreprise cliente exige que les lots expédiés par Simtech contiennent au plus 2% de défectueux. Les lots sont habituellement constitués de 5000 tubes de verre. Avant d'expédier les lots, Simtech effectue un contrôle en prélevant au hasard 200 tubes.

1. En utilisant un risque de α =0.05 de rejeter à tort un lot dont la proportion de défectueux est de 2% (ou mieux), quelle est la valeur critique de la proportion de défectueux dans un échantillon de taille n=200 qui ne doit pas être dépassée pour considérer un lot comme acceptable ?
2. On doit expédier un lot de 5000 tubes. Lors du contrôle final, on a observé 4 tubes de verre défectueux dans un échantillon de 200 tubes. Est-ce que ce lot peut être considéré comme acceptable d'après les exigences de Giscom ?
3. Giscom réceptionne des lots de Simtech en prélevant également 200 tubes et fait usage de la même règle de décision que Simtech. Quelles sont les chances sur 100 d'accepter un lot comportant 6% de tubes défectueux alors que Simtech certifie 2% de défectueux dans 95% des cas ? Comment appelle-t-on ce risque ?

Cas 4:

Un fabriquant de fournitures électriques fabrique des résistances dont la valeur nominale doit être de 1000 ohms (ohm-mètre)). Pour vérifier le procédé de fabrication, on prélève un échantillon aléatoire de 64 résistances. On mesure ces résistances avec un ohmètre de précision et les calculs de la moyenne et l'écart-type conduisent aux valeurs suivantes :

$\bar{x}$ = 990 ohms, s= 100ohms.

1. En supposant que le risque de première espèce est fixé à 0.05, élaborer une règle de décision qui permettrait de tester l'hypothèse nulle selon laquelle le procédé est centré à 1000 ohms.
2. Est-ce que l'hypothèse de normalité de la population (l'ensemble des valeurs ohmiques de la fabrication) est requise pour l'élaboration de la règle de décision en a)
3. Selon les résultats de l'échantillon de 64 résistances, doit-on rejeter l'hypothèse selon laquelle le procédé est centré à 1000 ohms ?
4. Quelle est la probabilité de ne pas rejeter l'hypothèse nulle H_0 : m=1000 ohms si le procédé opère en réalité à m= 1050 ohms ? Comment appelle-t-on ce risque ?

dard.

Cas 5 :

La résistance d'un composant électronique doit être, en moyenne, de 400 ohms. Un échantillon de 16 composants, prélevé d'un grand lot, conduit aux résultats suivants :

392	396	386	389
388	387	403	397
401	391	400	402
394	406	406	400

On considère que la distribution de la résistance est celle d'une loi normale.

1. Peut-on considérer, au seuil de signification $\alpha = 0{,}05$, que le lot respecte la norme de 400 ohms ?

2. Avec les résultats de cet échantillon, calculer un intervalle de confiance ayant un niveau de confiance de 95% de contenir la vraie moyenne. Est-ce que cet intervalle contient le norme spécifiée ?

Cas 6 :

Une entreprise fabrique des dispositifs électroniques dont la durée de vie moyenne est de 800 heures. La durée de vie des dispositifs est distribuée normalement avec un écart type $\sigma = 50$ heures. Pour vérifier la qualité des dispositifs, un échantillon de 25 dispositifs est soumis à un essai de fiabilité et on adopte la règle de décision suivante :

Les dispositifs sont de qualité inacceptable si la durée de vie moyenne ~~est~~ de 25 dispositifs est inférieur à 783.55 heures ; on les considère de qualité acceptable si la durée de vie moyenne est supérieure ou égale à 783.55 heures. → la valeur critique

1. Quelles sont les hypothèses statistiques que l'on veut tester avec cette règle de décision ?

2. Quelle est la probabilité de rejeter à tort un lot de dispositifs de qualité acceptable ?

3. Quel est le risque de deuxième espèce pour chacune des valeurs suivantes de μ : 750, 760, 770, 780, 800.

4. Le responsable du contrôle a décidé de modifier son plan de contrôle en prélevant 36 dispositifs de la production (au lieu de 25). Quelle règle de décision doit il alors adopter pour tester les hypothèses spécifiques en 1)

26-10-2015

Tests d'hypothèses

Cas 1

$n = 45$

- H_0 : $m = 16$ "conformité"
- H_1 : $m \neq 16$ "Non conformité"
- $\alpha = 5\%$

T.C.L $\bar{X}_n \sim N(m ; \sigma/\sqrt{n})$

$1-\alpha = \mathbb{P}[m-\varepsilon < \bar{x} < m+\varepsilon]$
$= \mathbb{P}\left[-\frac{\varepsilon\sqrt{n}}{\sigma} < \frac{\bar{x}-m}{\sigma/\sqrt{n}} < \frac{\varepsilon\sqrt{n}}{\sigma}\right]$

$\Pi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) = 1-\alpha/2 = 0,975$

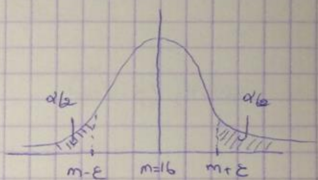
D'après la table $N(0,1)$: $\frac{\varepsilon\sqrt{n}}{\sigma} = 1,96$ où $\sigma = 0,8$

A.N $\varepsilon = \frac{1,96 \times 0,8}{\sqrt{30}}$

$\varepsilon = 0,288$

$V_1^* = 15,71$

$V_2^* = 16,29$



Règle de rejet :

si $\bar{x} \in [15,71 ; 16,29]$, Ho accepté

si $\bar{x} \notin [15,71 ; 16,29]$; H_0 rejeté

Cas 2

$n = 45$

- Ho $m = 2,2$ "Conformité"
- H_1 $m \neq 2,2$ "Non conformité"
- $\alpha = 2\% = 0,02$

T.C.L $\bar{X}_n \sim N(m, \sigma/\sqrt{n})$

$1-\alpha = \mathbb{P}[m-\varepsilon < \bar{x} < m+\varepsilon]$
$= \mathbb{P}\left[-\varepsilon\sqrt{n}/\sigma < \frac{\bar{x}-m}{\sigma/\sqrt{n}} < \frac{\varepsilon\sqrt{n}}{\sigma}\right)$

$\Pi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) = 1-\alpha/2 = 0,99$

D'après la table $N(0,1)$:

$\frac{\varepsilon\sqrt{n}}{\sigma} \approx 2,33$ où $\sigma = S = 0,2$

A.N $\varepsilon = \frac{2,33 \times 0,2}{\sqrt{45}} = 0,069$.

$V_1^* = 2,13$

$V_2^* = 2,27$

Cas 3

1) $n = 200$

- Ho : $p = 2\% = 0,02$ "Conformité"
- H_1 : $p > 0,02$ "Non conformité"
- $\alpha = 5\%$

$V^* = p + \varepsilon$ $f = 2,03\%$

T.C.L $F_n \sim N\left(p ; \sqrt{p(1-p)/n}\right)$

$0,05 = 1-\alpha = \mathbb{P}[F_n < V^*] = \mathbb{P}\left[\frac{F_n - p}{\sqrt{\frac{p(1-p)}{n}}}\right] < \varepsilon\sqrt{\frac{n}{p(1-p)}}\ ]$

$= \Pi\left(\varepsilon\sqrt{\frac{n}{p(1-p)}}\right) = 0,95$

D'après la table $N(0,1)$:

$$\varepsilon \sqrt{\frac{n}{p(1-p)}} \simeq 1,65$$

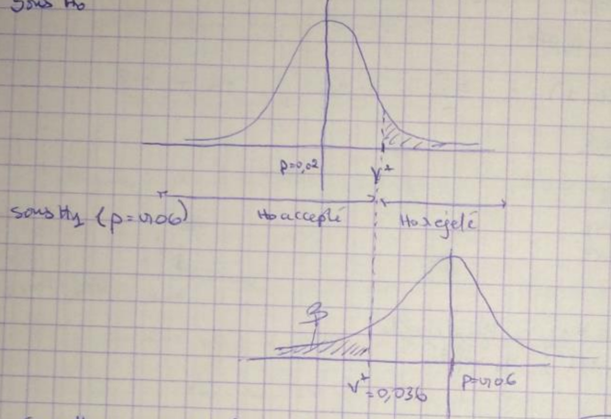
$$\varepsilon = 1,65 \sqrt{\frac{p(1-p)}{n}}$$

A.N $\varepsilon = 0,016$

$$v^* = 0,02 + 0,016$$

$$v^* = 0,036 = 3,6\,\%$$

3) Sous H_0



Sous H_1 $p = 0,06$ et $F_n \sim N\left(p \,;\, \sqrt{\frac{p(1-p)}{n}}\right)$

$$\beta = \mathbb{P}\left[F_n < 0,036\right]$$

$$= \mathbb{P}\left[\frac{F_n - 0,06}{0,0168} < \frac{0,036 - 0,06}{0,168}\right]$$

$$= \Pi(-1,43)$$

$$= 1 - \Pi(1,43)$$

$$= 1 - 0,9236$$

$$= 0,0764 \simeq 7,64\,\%$$

cas 4

$n = 64$

H_0 : $m = 1000$

H_1 : $m \neq 1000$

$\alpha = 5\%$

T.C.L $\bar{x} \sim N(m, \sigma/\sqrt{n})$, $\sigma \simeq S = 100$

$$\Pi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) = 1 - \alpha/2 \quad \text{où } \alpha = 0,05$$

D'après la table $N(0,1)$

$$\frac{\varepsilon\sqrt{n}}{\sigma} = 1,96$$

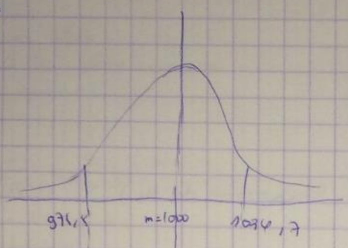
$$\varepsilon = \frac{196}{8} =$$

$v_1^* = 975,5$

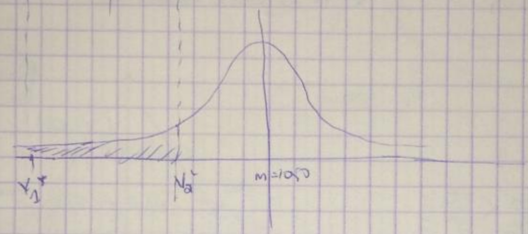
$v_2^* = 1024,5$

$990 \in [v_1^*, v_2^*]$

Sous H_0



Sous H_1



Sous H_1

$$X_n \sim N(1050 ; \frac{100}{8})$$

$$\beta = \mathbb{P}[975,5 < \bar{X}_n < 1024,5]$$

$$= \Pi(-2,04) - \Pi(-5,96)$$

$$= \Pi(2,04) + \Pi(5,96)$$

$$= 1 - \Pi(2,04) - (1 - \underbrace{\Pi(5,96)}_{\simeq 0})$$

$$= 1 - \Pi(2,04)$$

$$= 1 - 0,97933$$

$$\beta = 0,02067$$

$$\boxed{\beta \simeq 0,02}$$ 2%

Cas 5 02-10-2015

1) X "Résistance d'un composant électronique"
$X \sim N(m, \sigma)$ où σ inconnu

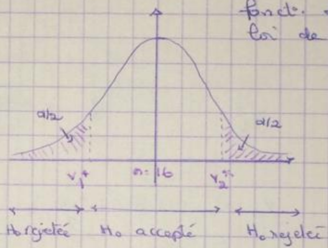
* $n = 16$
* H_0 : "$m = 400$" conformité
* H_1 : "$m \neq 400$" non conformité.
* $\alpha = 5\ \%$.

Cherchons les 2 valeurs critiques (d'après l'expression de H_1) V_1^* et V_2^*.
$X_i \sim N(m, \sigma)$ avec $X_{i=1,\dots,16}$

$$\frac{\bar{x} - m}{S^*/\sqrt{n}} \sim T_{(n-1)}$$ $\bar{x}$: la moyenne empirique.

$V_1^* = m - \varepsilon$
$V_2^* = m + \varepsilon$

fonct. densité de la loi de $\bar{X}$



$$1 - \alpha = \mathbb{P}\,[V_1^* < \bar{x} < V_2^*]$$

$$1 - 0{,}05 = \mathbb{P}\left[\frac{-\varepsilon\sqrt{n}}{S^*} < \frac{\bar{x} - m}{S^*/\sqrt{n}} < \frac{\varepsilon\sqrt{n}}{S^*}\right]$$

$$t_\alpha = \frac{\varepsilon\sqrt{n}}{S^*}$$

Table de Student :
Ds la première ligne on lit la valeur de α.
Ds la " colonne " " " de $n-1$
→ l'intersect° des 2 valeurs est t_α.

$$\varepsilon = \frac{t_\alpha S^*}{\sqrt{n}}$$ $t_\alpha = 2{,}1315$

Rappel
$$S^* = \left(\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})\right)^2$$

2) Estimation avec un intervalle de confiance avec σ inconnu.
a, b ? $\mathbb{P}\,[a < m < b] = 0{,}95$

$a = \bar{x} - t_\alpha\, S^*/\sqrt{n}$ $b = \bar{x} + t_\alpha\, S^*/\sqrt{n}$

où $t_\alpha = 2{,}1315$

$$\mathbb{P}\left[-t_\alpha < \frac{\bar{x} - m}{S^*/\sqrt{n}} < t_\alpha\right] = 0{,}95$$

Cas 6

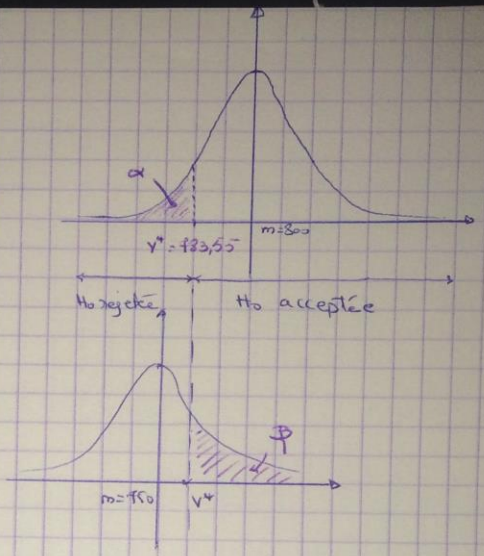
1) Tests d'hypothèses :
* $n = 25$
* H_0 : $m = 800$ "Conformité"
* H_1 : $m < 800$ "Non conformité"
* α Risque de 1re espèce.

X "Durée de vie d'un dispositif"
$X \sim N(m, \sigma)$ où $\sigma = 50$

$\sigma = 50$
$n = 25$

2)

$\alpha = \mathbb{P}[\bar{x} < V^*]$



$\bar{x} \sim N(m, \sigma/\sqrt{n})$
parce que le phénomène de départ est normal.

$\alpha = \mathbb{P}\left[\frac{\bar{x}-m}{\sigma/\sqrt{n}} < \frac{783,55-800}{10}\right]$

$= \Pi(-1,645)$

$= 1 - \Pi(1,645)$

$= 1 - 95\,053$

$= 0,04947$

$= 4,947\ \%$

3)

$\beta = \mathbb{P}[\bar{x} > 783,55]$
où $\bar{x} \sim N(750, 10)$

$\beta = \mathbb{P}\left[\frac{\bar{x}-750}{10} > 3,34\right]$

$= 1 - \Pi(3,34)$

$= 1 - 0,04947$

$= 4,947\%$

→ Même Démarche pour les autres valeurs.

le seul changement est au niveau de la taille de l'échantillon.
⇒ la valeur critique sera un peu plus proche de la norme.

$\alpha = \mathbb{P}(\bar{x} < m - \varepsilon)$

$= \mathbb{P}\left(\frac{\bar{x}-800}{10} < -\varepsilon/10\right]$

$= 1 - \Pi(\varepsilon/10)$

$= 1 - 0,95$ (D'après la table)

$= 0,05$

$\varepsilon/10 = 1,65$
$\varepsilon = 16,5$

Donc $V^* = 800 - \varepsilon = 783,5$