Test d'hypothèse

Ex 1: 1. Quelle est la probabilité d'obtenir entre 40 et 60 faces au cours de 100 jets d'une pièce de monnaie?
2. Pour vérifier l'hypothèse qu'une pièce est honnête, on adopte la règle de décision suivante:
(1)l'hypothèse est acceptée si le nombre de faces dans un échantillon unique est compris entre 40 et 60 inclusivement au cours des 100 jets,
(2) l'hypothèse est rejetée dans tous les autres cas.
a) Quelle est la probabilité de rejeter l'hypothèse alors qu'elle serait tout à fait convenable?
b) Interpréter graphiquement la règle de décision et le résultat de **a)**.
c) Quelles conclusions peut-on tirer si l'échantillon présente 53 faces? 60 faces?
3. Enoncer une règle de décision pour tester que l'hypothèse qu'une pièce est honnête si un échantillon de 64 jets est tiré et que le niveau de signification est (a) *0,05* et (b) *0,01*.

1. D'après la distribution binômiale, la probabilité cherchée est :

$$C_{100}^{40}\left(\frac{1}{2}\right)^{40}\left(\frac{1}{2}\right)^{60} + C_{100}^{41}\left(\frac{1}{2}\right)^{41}\left(\frac{1}{2}\right)^{59} + \ldots + C_{100}^{60}\left(\frac{1}{2}\right)^{60}\left(\frac{1}{2}\right)^{40}$$

La moyenne et l'écart-type du nombre de faces obtenues en 100 jets sont :

$$\mu = np = 100\left(\frac{1}{2}\right) = 50 \qquad \sigma = \sqrt{npq} = \sqrt{(100)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)} = 5$$

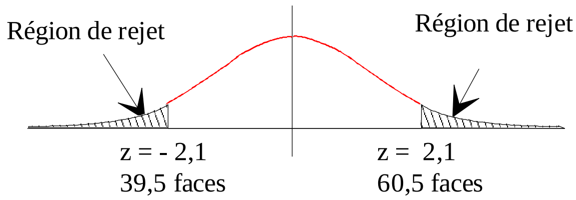
Comme *np* et *nq* sont tous deux supérieurs à 5, l'approximation normale à la distribution binômiale est valide pour l'évaluation de la somme ci-dessus.
D'un point de vue continu, l'intervalle fermé 40-60 est équivalent à [39,5; 60,5]

En unités réduites: 39,5 correspond à $\frac{39{,}5-50}{5} = -2{,}10$ et 60,5 correspond à $\frac{60{,}5-50}{5} = 2{,}10$

Probabilité cherchée = aire sous la courbe normale comprise entre $z = -2{,}10$ et $z = 2{,}10$
= 2(aire entre $z = 0$ et $z = 2{,}10$) = $2\times0{,}4821 = 0{,}9642$

2. a) D'après le **1.** la probabilité de ne pas obtenir entre 40 et 60 faces si la pièce est honnête est $1 - 0{,}9642 = 0{,}0358$. Il en résulte que la probabilité de rejeter cette hypothèse alors qu'elle est valide est égale à 0,0358.

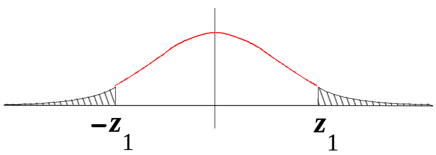
b) La règle de décision est illustrée graphiquement par la figure qui représente la distribution de probabilité des faces pour 100 jets d'une pièce honnête.
Si un échantillon unique de 100 jets donne un résultat z, compris entre –2,10 et 2,10, nous accepterons l'hypothèse; celle-ci sera rejetée dans tous les autres cas et nous en concluerons que la pièce est truquée.



L'erreur commise en rejetant l'hypothèse alors qu'elle devrait être acceptée est une erreur de première espèce de la règle de décision et, d'après a, sa probabilité est égale à 0,0358, qui est représenté par l'aire ombrée totale de la figure.
Si un échantillon unique de 100 jets conduit à un résultat z contenu dans l'aire ombrée, nous pouvons dire que ce résultat z est significativement différent de ce que l'on pourrait espérer si l'hypothèse était juste. Pour cette raison, l'aire totale ombrée (c'est à dire la probabilité d'une erreur de première espèce) est appelée niveau de signification de la règle de décision, égal à 0,0358 dans le cas.
Nous parlerons donc du rejet de l'hypothèse à un niveau de signification de 0,0358 ou 3,58%.

c) D'après la règle de décision, nous devrions accepter l'hypothèse de l'honnêteté de la pièce dans les deux cas. On pourrait cependant discuter le fait que si une face de plus seulement avait été retournée, l'hypothèse aurait dû être rejetée; c'est le cas chaque fois qu'il faut trancher.

3. <u>Première méthode :</u> Si le niveau de signification est 0,05 alors
$z_1 = 1{,}96$ Une règle de décision possible est alors :
Accepter l'hypothèse que la pièce est honnête si Z est compris entre –1,96 et 1,96.
Rejeter l'hypothèse dans tous les autres cas.



Pour exprimer numériquement la règle de décision en fonction du nombre de faces obtenues en 64 jets, notons que la moyenne et l'écart-type de la distribution binômiale stricte des faces sont :

$\mu = np = 64(0,5) = 32$ et $\sigma = \sqrt{npq} = \sqrt{64(0,5)(0,5)} = 4$

Dans l'hypothèse que la pièce est honnête. Alors $Z = \dfrac{X-\mu}{\sigma} = \dfrac{X-32}{4}$

Si $Z = 1,96$, $(X - 32) / 4 = 1,96$ ou $X = 39,84$.
Si $Z = -1,96$, $(X - 32) / 4 = -1,96$ ou $X = 24,16$.
La règle de décision devient

Accepter l'hypothèse si le nombre de faces est compris entre 24,16 et 39,84, c'est à dire entre 24 et 39, inclus.
Rejeter l'hypothèse dans tous les autres cas.

Seconde méthode : Avec une probabilité de 0,95, le nombre de faces sera compris entre $\mu - 1,06\,\sigma$ et $\mu + 1,96\,\sigma$

soit $np - 1,96\sqrt{npq}$ et $np + 1,96\sqrt{npq}$ ou $32 - 1,96\times4 = 24,16$ et $32 + 1,96\times4 = 39,84$,

ce qui conduit à la règle de décision énoncée ci-dessus.

Troisième méthode : $-1,96 < Z < 1,96$ est équivalent à $-1,96 < (X - 32) / 4 < 1,96$.
Par conséquent, $-1,96\times4 < (X - 32) < 1,96\times4$ ou $32 - 1,96\times4 < X < 32 + 1,96\times4$,
c'est à dire $24,16 < X < 39,84$,qui conduit également à la règle de décision ci-dessus.

b) Si le niveau de signification est 0,01, $z_1 = 2,58$ (plus précisément 2,575)
En suivant la seconde méthode ci-dessus, nous voyons qu'avec une probabilité 0,99 le nombre de faces sera compris entre $\mu - 2,58\,\sigma$ et $\mu + 2,58\,\sigma$, c'est à dire $32 - 2,58\times4 = 21,68$ et $32 + 2,58\times4 = 42,32$.
La règle de décision devient:

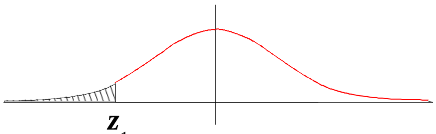
Accepter l'hypothèse si le nombre de faces est compris entre 22 et 42 inclus.
Rejeter l'hypothèse dans tous les autres cas.

Ex 2 : Le fabricant d'un médicament annonce qu'un de ses produits est efficace à 90%, en supprimant une allergie dans un délai de 8 heures. Dans un échantillon de 200 personnes, le résultat a été effectif pour 160 d'entre elles. Déterminer si l'affirmation du fabricant est légitime au seuil de 1%.

Soit p la probabilité du soulagement par utilisation du médicament. Nous aurons alors à choisir entre deux hypothèses:

H_0: $p = 0,9$, et l'affirmation est correcte.
H_1: $p < 0,9$ et l'affirmation est fausse

Nous choisissons un test unilatéral car nous sommes intéressés à déterminer si la proportion des individus soulagés n'est pas trop basse



z_1

Si nous choisissons un niveau de signification de 0,01, c'est à dire si l'aire hachurée sur la figure est 0,01, $z_1 = -2,33$ comme on peut le voir d'après le problème par symétrie. La règle de décision sera:

L'affirmation est fausse si Z est inférieur à $-2,33$ (en ce cas, nous rejetons H_0)
Dans tous les autres cas, l'affirmation est légitime (et nous acceptons H_0).

Si H_0 est vraie, $\mu = np = 200\times0,9 = 180$ et $\sigma = \sqrt{npq} = \sqrt{(200)(0,9)(0,1)} = 4,23$

En unités réduite 160 vaut $(160 - 180)/4,23 = -4,73$, très inférieur à $-2,33$. Alors, d'après notre règle de décision, nous concluerons que l'affirmation est fausse et que les résultats obtenus sur l'échantillon sont hautement significatifs.

Ex 3 : 1. La durée de vie moyenne d'un échantillon de 100 tubes fluorescents a été établie par le calcul à 1570 heures avec un écart-type de 120 heures. Si μ est la moyenne de la population, vérifier l'hypothèse $\mu = 1600$ heures relativement à l'hypothèse $\mu \neq 1600$ heures, avec un niveau de signification de 0,05 puis de 0,01.
2. Résoudre alors le problème dans le cas où la deuxième hypothèse devient $\mu < 1600$ heures.

1. Nous devons choisir entre les deux hypothèses H_0 : $\mu = 1600$ heures et H_1 : $\mu \neq 1600$ heures
Nous utiliserons un test bilatéral car $\mu \neq 1600$ comprend des valeurs aussi bien inférieures que supérieures à 1600.
a) Pour un test bilatéral au niveau de 0,05, la règle de décision est la suivante :

Rejeter H_0 si le résultat z est à l'extérieur de l'intervalle $-1,96$ à $1,96$
Accepter H_0 (ou suspendre toute décision) dans les autres cas.

La statistique en jeu est la moyenne $\overline{X}$ de l'échantillon. La distribution d'échantillonnage de $\overline{X}$ a une moyenne $\mu_{\overline{X}} = \mu$

et un écart-type $\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}}$, où μ et σ sont la moyenne et l'écart-type de la population comprenant tous les tubes de même origine.

Dans l'hypothèse H_0, nous avons $\mu = 1600$ et $\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{120}{\sqrt{100}} = 12$

En utilisant l'écart-type de l'échantillon comme estimation de σ. Puisque $Z = \frac{(\bar{X} - 1600)}{12} = \frac{(2570 - 1600)}{12} = -2{,}50$ est à l'extérieur de l'intervalle [– 1,96; 1,96], nous rejetons H_0 au niveau de signification 0,05.

b) Si le niveau de signification est 0,01, l'intervalle précédent devient [- 2,58; 2,58]. Comme le résultat z = 2,50 est à l'intérieur de l'intervalle, nous acceptons H_0 (ou nous suspendons toute décision) au niveau de signification 0,01.

2. Nous devons choisir entre les deux hypothèses $H_0 : \mu = 1600$ heures $H_1 : \mu < 1600$ heures
et nous utiliserons ici un test unilatéral
a) Si le niveau est 0,05, la région hachurée sur la figure vaut 0,05 et il résulte que $z_1 = -1{,}645$.
La règle de décision est :

Rejeter H_0 si $Z < -1{,}645$
Accepter H_0 (ou suspendre toute décision) dans tous les autres cas.

Puisque, comme dans le **1.** $z = -2{,}50 < -1{,}645$, nous rejetons H_0 à un niveau de signification de 0,05.
Notons que le résultat est le même que celui que nous avons trouvé en utilisant un test bilatéral.

b) Si le niveau de signification est 0,01; $z_1 = -2{,}33$; d'où la règle de décision:

Rejeter H_0 si $Z < -2{,}33$
Accepter H_0 (ou suspendre la décision) dans tous les autres cas.

Puisque, comme dans le problème précédent, $z = -2{,}50 < -2{,}33$, nous rejetons H_0 à un niveau de signification de 0,01.
Notons que nous aboutissons à une décision différente de la décision atteinte au problème précédent. Ce n'est pas étonnant puisque nous testions H_0 contre une hypothèse différente dans chaque cas.

Ex 4 : Le poids moyen de 50 étudiants d'une université donnée et, qui se sont montrés très amateurs de lutte sportive, s'est trouvé être de 68,2 kg, avec un écart-type de 3,6 kg, tandis qu'un autre groupe de 50 étudiants nettement moins intéressés par le sport, ont présenté un poids moyen de 67,5 kg, avec un écart-type de 2,8 kg. Tester l'hypothèse selon laquelle les étudiants amateurs de sport pèsent plus que les autres étudiants.

Nous devons choisir entre les hypothèses :

H_0 : $\mu_1 = \mu_2$, et il n'y a pas de différence entre les poids moyens.
H_1 : $\mu_1 > \mu_2$, et le poids moyen du premier groupe est supérieur à celui du second groupe.

Dans l'hypothèse H_0, $\mu_{\bar{X}_1 - \bar{X}_2} = 0$ $\quad \sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{(2{,}5)^2}{50} + \frac{(2{,}8)^2}{50}} = 0{,}53$

où nous avons utilisé l'écart-type de l'échantillon comme estimateur de σ_1 et σ_2.

Alors $Z = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{\bar{X}_1 - \bar{X}_2}} = \frac{68{,}2 - 67{,}5}{0{,}53} = 1{,}32$

Sur la base d'un test unilatéral à un niveau 0,05, nous rejetterions H_0 si z était supérieur à 1,645.Ce n'est pas le cas; aussi nous ne pouvons rejeter cette hypothèse au niveau de signification 0,05. Remarquons, néanmoins, que cette hypothèse peut être rejetée à un niveau de 0,10 si nous acceptons de la faire avec le risque de nous tromper sur une probabilité de 0,1, c'est à dire avec 1 chance sur 10.

Ex 5: Dans le passé, l'écart-type des poids de colis remplis à 40,0 kg par une machine donnée était de 0,25 kg. Un échantillon aléatoire de 20 colis a présenté un écart-type de 0,32 kg. Est-ce que cette augmentation apparente de dispersion est significative au niveau de a) 0,05 et b) 0,01?

Nous devons choisir entre les hypothèses :

H_0 : $\sigma = 0{,}25$ kg, et le résultat est dû au hasard.
H_1 : $\sigma > 0{,}25$ kg, et la dispersion a augmenté.

La valeur de χ^2 pour l'échantillon est $\chi^2 = \frac{ns^2}{\sigma^2} = \frac{20(0{,}32)^2}{(0{,}25)^2} = 32{,}8$

a) Test unilatéral: si $\chi^2 > \chi^2_{0,95}$; c'est à dire $\chi^2 > 30{,}1$ pour $\nu = 20 - 1 = 10$ degrés de liberté, H_0 doit être rejetée au niveau de signification 0,05, ce qui est le cas ici.
b) Test unilatéral: si $\chi^2 > \chi^2_{0,99}$, c'est à dire $\chi^2 > 36{,}2$ pour 19 degrés de liberté, H_0 devrait être rejetée au niveau de signification 0,01, ce qui n'est pas le cas ici.
Nous en concluons que la dispersion a problament augmenté. Il y a lieu d'opérer une révision de la machine.

Ex 6 : 1. La longueur des pièces (test bilatéral)
Dans un atelier une machine fabrique des pièces en grande série; on s'intéresse à leur longueur mesurée en cm. On admet que la variable aléatoire X qui, à chaque pièce tirée au hasard dans la production associe sa longueur, suit une loi normale de moyenne m et d'écart-type $\sigma = 0{,}14$.
afin de contrôler le fait que la moyenne m des longueurs des pièces produites est 150, on se propose de construire un test d'hypothèse.
On prélève des échantillons aléatoires de 49 pièces (chaque échantillon étant obtenu par tirage avec remise). A chaque échantillon ainsi défini, on associe la moyenne $\bar{x}$ des longueurs des 49 pièces; on définit ainsi une variable aléatoire $\bar{X}$.
L'hypothèse nulle est H_0: $m = 150$. L'hypothèse alternative est H_1: $m \neq 150$. Le seuil de signification du test est fixé à 0,05.
a) Quelle est, sous l'hypothèse nulle H_0, la loi de la variable aléatoire $\bar{X}$.
Déterminer le nombre réel positif h tel que $P(150 - h \leq \bar{X} \leq 150 = h) = 0{,}95$.
b) Enoncer la règle de décision permettant d'utiliser ce test.
c) La moyenne observée sur un échantillon de 49 pièces est $\bar{x} = 149{,}9$.
Que peut-on conclure au seuil de signification 5% quant à la qualité des pièces produites?

2. Problème de longueur (suite) (test bilatéral)
Des pièces sont fabriquées en grande série. On s'intéresse à leur longueur. On suppose que la variable aléatoire X qui, à chaque pièce prélevée au hasard dans la production, associe sa longueur exprimée en millimètres, suit la loi normale de moyenne $m = 14$ et d'écart-type $\sigma = 0{,}3$.
Dans un lot important de pièces reçues, un client prélève un échantillon de 100 pièces. On admet que l'écart-type du lot est celui de la production. On désigne par $\bar{X}$ la variable aléatoire qui, à tout échantillon aléatoire de 100 pièces, associe la moyenne des longueurs des 100 pièces. (On peut assimiler ces prélèvements à des tirages avec remise).
a) On admet que $\bar{X}$ suit un loi normale; en préciser les paramètres.
b) Déterminer le nombre réel positif h, tel que la probabilité que la moyenne des longueurs appartienne à l'intervalle $[14–h, 14+h]$ soit 0,95.
c) Le constructeur affirme que la moyenne des longueurs est 14 mm. Pour vérifier cette affirmation, le client mesure les longueurs des pièces d'un échantillon, puis calcule la moyenne des 100 mesures. Il trouve 13,93. Peut-on, au risque de 5% admettre que la moyenne des longueurs du lot est 14 mm?

1. a) Sous l'hypothèse H_0 , $\bar{X}$ suit la loi $\mathcal{N}$ (150; 0,02) : $\dfrac{\sigma}{\sqrt{n}} = \dfrac{0{,}14}{7} = 0{,}02$

$\text{Prob}(150 - h \leq \bar{X} \leq 150 = h) = 0{,}95$ correspond à Prob($-t \leq z \leq t$) = 0,95 donc à $t = 1{,}96$.
d'où $h = 1{,}96 \times 0{,}02 = 0{,}0394 \approx 0{,}04$. On trouve alors Prob ($149{,}96 \leq \bar{X} \leq 150{,}04$) = 0,95.

b) Règle de décision :
Si la moyenne d'un échantillon de 49 pièces est comprise entre 149,96 et 150,04 on peut considérer, au seuil de 5%, que la moyenne des longueurs des pièces est de 150. Sinon, la moyenne est différente de 150.

c) La moyenne observée sur un échantillon de 49 pièces est de 149,9.
Au seuil de 5%, on peut considérer que les pièces ne sont pas de bonne qualité.

rem : erreur de 1ère espèce, il y a 5% de chance de refuser des pièces de mauvaise qualité.

2. D'après les données, le lot a une moyenne de 14 et un écart-type de 0,3.

a) $\bar{X}$ suit donc la loi $\mathcal{N}$ (14; 0,03) : $\dfrac{\sigma}{\sqrt{n}} = \dfrac{0{,}3}{\sqrt{100}} = 0{,}03$.

b) Prob ($14 - h \leq \bar{X} \leq 14 + h$) = 0,95 correspond à Prob($-t \leq z \leq t$) = 0,95 donc à $t = 1{,}96$.
d'où $h = 1{,}96 \times 0{,}03 = 0{,}0588 \approx 0{,}06$. On trouve alors Prob ($13{,}94 \leq \bar{X} \leq 14{,}06$) = 0,95.

c) D'après le **b)** la règle de décision au seuil de 5% s'enonce :
Si la moyenne d'un échantillon de 100 pièces est comprise entre 13,94 et 14,06 alors la moyenne des longueurs du lot peut être considérée comme égale à 14 mm; sinon la moyenne est différente de 14 mm.

La moyenne observée de 13,93 n'est pas dans l'intervalle, donc au seuil de 5%, la moyenne des longueurs du lot n'est pas de 14 mm.

rem : erreur de première espèce, il y a 5% de chance de refuser des pièces de bonne qualité.

Ex 7 : Dosage et qualité

On a contrôlé le dosage d'un produit dans un mélange à la sortie d'une chaîne de conditionnement. Pour un échantillon de 100 lots (tirés au hasard et avec remise) de 5 kilogrammes de mélange analysés, on a obtenu les résultats suivants où P_i représente la masse du produit exprimée en grammes et n_i l'effectif correspondant:

p_i	[141;143[	[143;145[	[145;147[	[147;149[	[149;151[	[151;153[	[153;155[	[155;157[	[157;159[	[159;161]
n_i	1	5	6	21	32	22	7	4	1	1

Un lot de 5 kilogrammes de mélange est dit de "qualité supérieure" s'il contient entre 147 grammes et 155 grammes de produit.

1. Déterminer pour cet échantillon de taille n = 100 le pourcentage f de lots de qualité supérieure.

2. Soit F la variable aléatoire qui, à tout échantillon de taille n = 100 prélevé au hasard et avec remise parmi l'ensemble des lots sortant de la chaîne de conditionnement, associe le pourcentage de lots de qualité supérieure de l'échantillon .On suppose que F suit la loi normale $\mathcal{N}\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$, où p est le pourcentage inconnu de lots de qualité extra dans l'ensemble des lots sortant de la chaîne de conditionnement.

Déterminer une estimation de p par un intervalle de confiance centré en f, avec le coefficient de confiance 98%.

3. Construire un test permettant d'accepter ou de rejeter au seuil de signification de 5% , l'hypothèse selon laquelle 90% des lots sortant de la chaîne de conditionnement sont de qualité supérieure.

4. Même question avec le seuil de signification de 1%.

5. Construire un test permettant d'accepter ou de rejeter au seuil de signification de 1%, l'hypothèse selon laquelle moins de 90% des lots sortant de la chaîne de conditionnement sont de qualité supérieure.

1. La somme des lots pour lesquels P_i est compris entre 147 et 155 est de 82 donc f = 82%.

2. On suppose alors que F suit la loi $\mathcal{N}\left(0,82; \sqrt{\frac{0,82 \times 0,18}{100}}\right)$.

Pour l'estimation par intervalle de confiance centré, on cherche alors a tel que Prob($0,82 - a \le F \le 0,82 + a$) = 0,98.

On a donc t = 2,33 et $a = 2,33 \times \sqrt{\frac{0,82 \times 0,18}{100}} \approx 0,0895$.

L'intervalle de confiance, avec le coefficient de confiance 98%, est [73,05; 90,95].

3. Test au seuil de 5% de l'hypothèse selon laquelle 90% des lots sont de qualité supérieure.

H_0 : " p = 0,9 " . Sous H_0 , F suit la loi $\mathcal{N}$ (0,9; 0,03) $\quad \sqrt{\frac{0,9 \times 0,1}{100}} = 0,03$.

Au seuil de 5%, t = 1,96 donc a = 1,96×0,03 = 0,0588. l'intervalle obtenu est : [84,12; 95,88]

règle de décision : Si la moyenne du lot est comprise entre 84,12 et 95,88 alors le lot est considéré comme étant de qualité supérieure. Sinon le lot n'est pas de qualité supérieure.

Pour l'échantillon tiré f = 0,82 donc il n'est pas considéré de qualité supérieure.

4. Test au seuil de 1% de l'hypothèse selon laquelle 90% des lots sont de qualité supérieure.

H_0 : " p = 0,9 " . Sous H_0 , F suit la loi $\mathcal{N}$ (0,9; 0,03) $\quad \sqrt{\frac{0,9 \times 0,1}{100}} = 0,03$.

Au seuil de 1%, t = 2,58 donc a = 2,58×0,03 = 0,0774. l'intervalle obtenu est : [82,26 ; 97,74]

règle de décision : Si la moyenne du lot est comprise entre 82,26 et 97,74 alors le lot est considéré comme étant de qualité supérieure. Sinon le lot n'est pas de qualité supérieure.

Pour l'échantillon tiré f = 0,82 donc il n'est pas considéré de qualité supérieure.

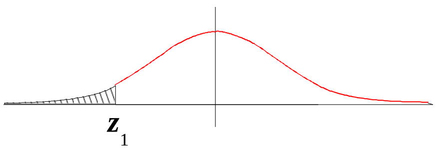
5. Test au seuil de 1% de l'hypothèse selon laquelle moins de 90% des lots sont de qualité supérieure.

H_0 : " p = 0,9 " ; H_1 " p < 0,9 ". Sous H_0 , F suit la loi $\mathcal{N}$ (0,9; 0,03) $\quad \sqrt{\frac{0,9 \times 0,1}{100}} = 0,03$.

Nous prenons un test unilatéral pour vérifier l'hypothèse H_1 :

Au seuil de 1%, t = 2,33 donc a = 2,33×0,03 = 0,0699.

donc Prob(83,01 ≤ F) = 0,99.



règle de décision : Si la moyenne du lot est inférieure à 83,01 alors H_1 est vraie et moins de 90% des lots sont de qualité supérieure, sinon il y a plus de 90% des lots qui sont de qualité supérieure.

Pour l'échantillon tiré f = 0,82 donc il y a moins de 90% des lots qui sont de qualité supérieure.

Ex 8 : On roule pour vous

Une entreprise utilise des camions pour transporter sa production. Elle dispose de 100 camions. Elle repère sur un échantillon de 30 jours choisis au hasard le nombre de camions en panne. Voici les résultats:

5 5 6 4 6 6 8 3 5 5 5 4 3 6 5 6 4 7 6 6 5 4 3 6 5 4 5 4 5 5

1. Calculer la moyenne $\bar{x}$ et l'écart-type σ du nombre de camions en panne chaque jour pour l'échantillon étudié.

2. A partir des résultats obtenus pour cet échantillon, proposer une estimation ponctuelle de la moyenne μ et de l'écart-type s du nombre de camions en panne chaque jour pour la population correspondant aux jours ouvrables de l'année.

3. On suppose que la variable aléatoire $\overline{X}$ qui, à tout échantillon de taille 30 prélevé au hasard et avec remise, associe la moyenne du nombre ce camions en panne chaque jour, suit la loi normale $\mathcal{N}\left(\mu, \frac{s}{\sqrt{n}}\right)$. On prend pour valeur de s l'estimation ponctuelle obtenue au **2.**

Déterminer un intervalle de confiance de la moyenne μ de la population avec le coefficient de confiance 95%.

4. Au garage où sont stationnés les camions, le responsable affirme qu'il y a, en moyenne, 4 camions en panne par jour. Un des chauffeurs prétend qu'il y en a 6.

Construire un test permettant d'accepter ou non, au seuil de signification de 5% et au vu des résultats sur 30 jours, le point de vue du responsable.

5. Même question pour le chauffeur

6. Reprendre les deux questions précédentes en utilisant le seuil de signification de 1%.

1. Un calcul simple donne $\bar{x} = \frac{151}{30} = 5{,}03$ et $\sigma = 1{,}1396 \approx 1{,}14$.

2. Estimation ponctuelle : $\mu = \frac{151}{30}$ et $s = \sqrt{\frac{30}{29}}\sigma \approx 1{,}16$.

On admet que $\overline{X}$ suit la loi $\mathcal{N}\left(\frac{151}{30}; \frac{1{,}16}{\sqrt{30}}\right)$

3. Intervalle de confiance avec le coefficient 95% : $t = 1{,}96$ donc $a = 1{,}96 \times \frac{1{,}16}{\sqrt{30}} \approx 0{,}4151$.

l'intervalle est alors [4,6185 ; 5,4481].

4. Au seuil de 5%. $t = 1{,}96$

Point de vue du responsable : H_0 " $\mu = 4$ " ; $\overline{X}$ suit la loi $\mathcal{N}\left(4; \frac{1{,}16}{\sqrt{30}}\right)$

$4 - 1{,}96 \times \frac{1{,}16}{\sqrt{30}} \approx 3{,}58$; $4 + 1{,}96 \times \frac{1{,}16}{\sqrt{30}} \approx 4{,}41$

règle de décision : Si la moyenne de l'échantillon est comprise entre 3,58 et 4,41 alors il y a en moyenne 4 camions en panne chaque jour.

Point de vue du chauffeur : H_0 " $\mu = 6$ " ; $\overline{X}$ suit la loi $\mathcal{N}\left(6; \frac{1{,}16}{\sqrt{30}}\right)$

$6 - 1{,}96 \times \frac{1{,}16}{\sqrt{30}} \approx 5{,}58$; $6 + 1{,}96 \times \frac{1{,}16}{\sqrt{30}} \approx 6{,}41$

règle de décision : Si la moyenne de l'échantillon est comprise entre 5,58 et 6,41 alors il y a en moyenne 6 camions en panne chaque jour.

5. Au seuil de 1%. $t = 2{,}58$

Point de vue du responsable : H_0 " $\mu = 4$ " ; $\overline{X}$ suit la loi $\mathcal{N}\left(4; \frac{1{,}16}{\sqrt{30}}\right)$

$4 - 2{,}58 \times \frac{1{,}16}{\sqrt{30}} \approx 3{,}45$; $4 + 1{,}96 \times \frac{1{,}16}{\sqrt{30}} \approx 4{,}54$

règle de décision : Si la moyenne de l'échantillon est comprise entre 3,45 et 4,54 alors il y a en moyenne 4 camions en panne chaque jour.

Point de vue du chauffeur : H_0 " $\mu = 6$ " ; $\overline{X}$ suit la loi $\mathcal{N}\left(6; \frac{1{,}16}{\sqrt{30}}\right)$

$6 - 2{,}58 \times \frac{1{,}16}{\sqrt{30}} \approx 5{,}45$; $6 + 2{,}58 \times \frac{1{,}16}{\sqrt{30}} \approx 6{,}54$

règle de décision : Si la moyenne de l'échantillon est comprise entre 5,45 et 6,54 alors il y a en moyenne 6 camions en panne chaque jour.

rem : aucun des test ne permet de trancher entre le patron et le chauffeur.