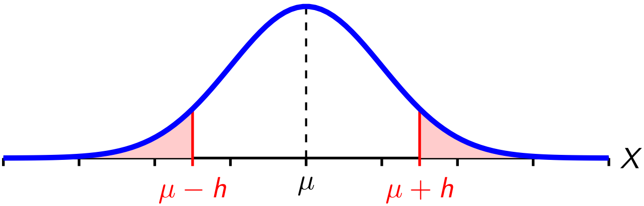
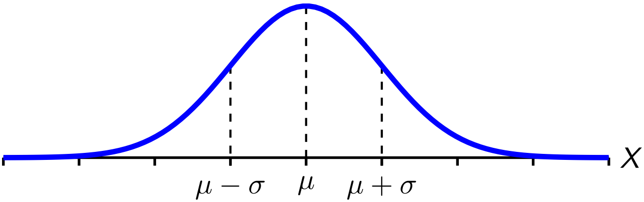


Statistiques inférentielles

I. Loi Normale centrée réduite

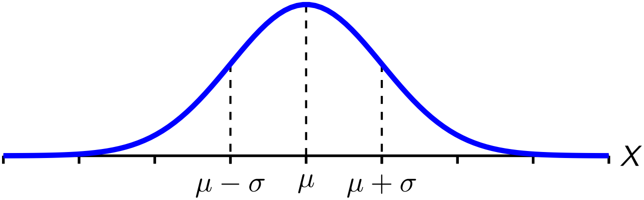
Besoin

X suit une loi normale $\mathcal{N}(\mu; \sigma^2)$.



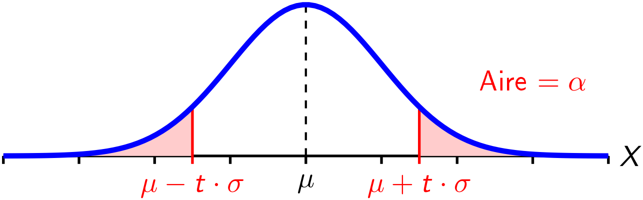
Besoin

X suit une loi normale $\mathcal{N}(\mu; \sigma^2)$.



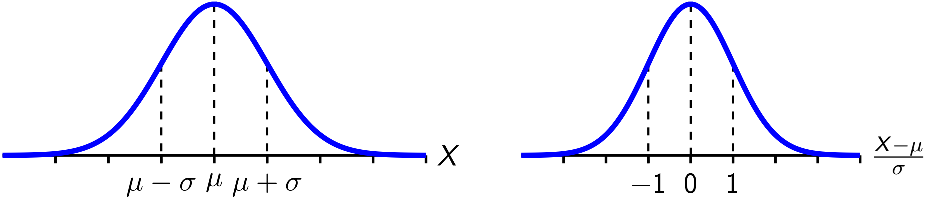
On se fixe α.

On cherche t tel que $p(X \notin [\mu - t \cdot \sigma; \mu + t \cdot \sigma]) = \alpha$



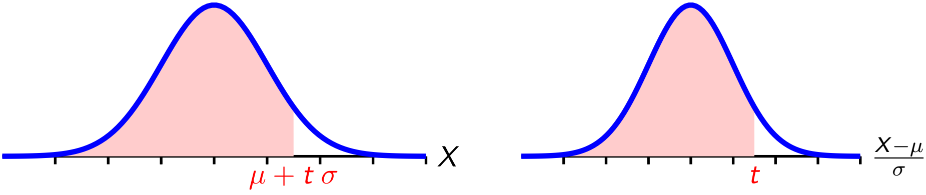
Autrement dit : $p(\mu - t \cdot \sigma \leqslant X \leqslant \mu + t \cdot \sigma) = 1 - \alpha$

$$X \text{ suit } \mathcal{N}(\mu;\sigma^2) \Rightarrow \frac{X-\mu}{\sigma} \text{ suit } \mathcal{N}(0;1^2)$$



Méthode 1 : Loi normale centrée réduite

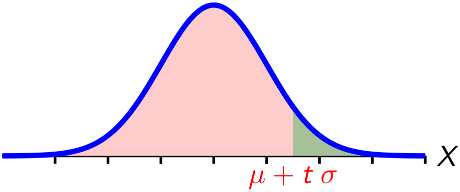
$$X \text{ suit } \mathcal{N}(\mu;\sigma^2) \Rightarrow \frac{X-\mu}{\sigma} \text{ suit } \mathcal{N}(0;1^2)$$

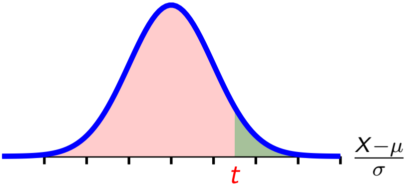


L'aire $\mathcal{A}$ est donnée par $\Pi(t)$
C'est une fonction connue.

Méthode 1 : Loi normale centrée réduite

$$X \text{ suit } \mathcal{N}(\mu; \sigma^2) \Rightarrow \frac{X-\mu}{\sigma} \text{ suit } \mathcal{N}(0; 1^2)$$

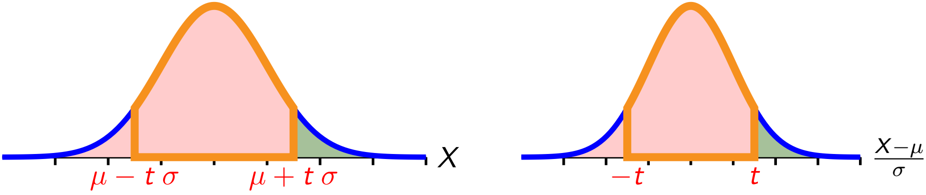




L'aire $\mathcal{A}$ est donnée par $\Pi(t)$
C'est une fonction connue.
L'aire a est donnée par $1 - \Pi(t)$

Méthode 1 : Loi normale centrée réduite

$$X \text{ suit } \mathcal{N}(\mu; \sigma^2) \Rightarrow \frac{X-\mu}{\sigma} \text{ suit } \mathcal{N}(0; 1^2)$$



L'aire $\mathcal{A}$ est donnée par $\Pi(t)$
C'est une fonction connue.
L'aire a est donnée par $1 - \Pi(t)$
L'aire $1 - \alpha$ est donc $1 - 2\,a = 1 - 2 \times [1 - \Pi(t)]$
$\Rightarrow 1 - \alpha = 2\Pi(t) - 1$

Exemple

X suit $\mathcal{N}(150; 4,5^2)$ et on veut $\alpha = 0,05 \Rightarrow 1 - \alpha = 0,95 = 95\,\%$.

X suit $\mathcal{N}(150; 4,5^2)$ et on veut $\alpha = 0,05 \Rightarrow 1 - \alpha = 0,95 = 95\,\%$.
Il faut trouver t tel que $2\Pi(t) - 1 = 1 - \alpha = 0,95 \Rightarrow \Pi(t) = 0,975$.
On cherche dans une table (extrait) :

t	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.961
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.969
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.975

Exemple

X suit $\mathcal{N}(150; 4,5^2)$ et on veut $\alpha = 0,05 \Rightarrow 1 - \alpha = 0,95 = 95\,\%$.
Il faut trouver t tel que $2\Pi(t) - 1 = 1 - \alpha = 0,95 \Rightarrow \Pi(t) = 0,975$.
On cherche dans une table (extrait) :

t	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.961
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.969
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.975

On trouve $t \simeq 1,96$ (ce qui est une valeur bien connue)

On sait donc que $p(150 - 1,96 \cdot 4,5 \leqslant X \leqslant 150 + 1,96 \cdot 4,5]) \simeq 0,95$

Méthode 2 : Calculatrice

Les calculatrices assez récentes disposent d'une fonction **Inverse Normale**. Il suffit de l'utiliser pour trouver que dans l'exemple précédent :

$$p(141,18 \leqslant X \leqslant 158,82) \simeq 0,95$$

Ce qu'il faut surtout retenir

1. $\alpha = 5\,\%$, il faut $t = 1,96$
2. $\alpha = 1\,\%$, il faut $t = 2,58$

II. Échantillonnage

On connaît la loi de X sur une population. On s'intéresse à un échantillon.

II. 1) Loi des moyennes

Dans une population de moyenne μ et d'écart-type σ.
On prélève un échantillon de taille n.
On relève la moyenne $\overline{X_n}$ de l'échantillon.

$\overline{X_n}$ suit approximativement une loi $\mathcal{N}$ de moyenne m d'écart-type $\dfrac{\sigma}{\sqrt{n}}$

Si la population suit une loi normale, ce n'est pas une approximation.

II. 1) Loi des moyennes

Dans une population de moyenne μ et d'écart-type σ.
On prélève un échantillon de taille n.
On relève la moyenne $\overline{X_n}$ de l'échantillon.

$$\overline{X_n} \text{ suit approximativement une loi } \mathcal{N} \text{ de moyenne } m \text{ d'écart-type } \frac{\sigma}{\sqrt{n}}$$

Si la population suit une loi normale, ce n'est pas une approximation.

Exemple : Une usine fabrique des tiges. En moyenne, le diamètre des tiges est de $20\,\mathrm{mm}$, avec un écart type de $0{,}1\,\mathrm{mm}$.

On prélève un échantillon de 100 tiges. Soit $\overline{X_{100}}$ la moyenne des diamètres de cet échantillon. Quelle est la probabilité que $19,99 \leqslant \overline{X_{100}} \leqslant 20,01$?

II. 1) Loi des moyennes

Dans une population de moyenne μ et d'écart-type σ.
On prélève un échantillon de taille n.
On relève la moyenne $\overline{X_n}$ de l'échantillon.

$$\overline{X_n} \text{ suit approximativement une loi } \mathcal{N} \text{ de moyenne } m \text{ d'écart-type } \frac{\sigma}{\sqrt{n}}$$

Si la population suit une loi normale, ce n'est pas une approximation.

Exemple : Une usine fabrique des tiges. En moyenne, le diamètre des tiges est de 20 mm, avec un écart type de 0,1 mm.

On prélève un échantillon de 100 tiges. Soit $\overline{X_{100}}$ la moyenne des diamètres de cet échantillon. Quelle est la probabilité que $19,99 \leqslant \overline{X_{100}} \leqslant 20,01$?

$\overline{X}$ suit la loi $\mathcal{N}$ de moyenne 20 et écart-type $\frac{0{,}1}{\sqrt{100}} = 0,01$. La probabilité recherchée est donc :

$$p(19,99 \leqslant \overline{X} \leqslant 20,01) \simeq 0,6826$$

Exercice 1

Une usine produit des disques en grande série. On appelle X la variable aléatoire qui à chaque disque, pris au hasard, associe son diamètre. On suppose que X suit la loi normale de moyenne $\mu = 50$ et d'écart-type $\sigma = 0,5$. On prélève au hasard des échantillons de 36 disques. On note Y la variable aléatoire qui pour chaque échantillon associe la moyenne des diamètres des 36 disques.

1. Quelle est la loi suivie par Y ?
2. Déterminez a pour que $p(\mu - a \leqslant Y \leqslant \mu + a) = 0,9$.

Exercice 2

Un traceur produit des cartouches ayant une forme de rectangle dont la longueur doit être de $10\,\mathrm{cm}$. Pour tester le réglage du traceur, on prélève 100 feuilles. Soit X la variable aléatoire qui pour chaque feuille associe la longueur du cartouche. On suppose que X suit la loi normale de moyenne $m = 10$ et d'écart-type $\sigma = 0,03$. On note $\overline{X_{100}}$ la variable qui à chaque échantillon associe la moyenne des longueurs.

1. Quelle est la loi suivie par $\overline{X_{100}}$?
2. Déterminez le réel positif h tel que $p\left(m - h \leqslant \overline{X_{100}} \leqslant m + h\right) = 0,95$.

Exercice 3

Une machine fabrique un très grand nombre de pièces, de longueur moyenne $m = 4{,}2\,\mathrm{mm}$ et d'écart-type $\sigma = 0{,}2\,\mathrm{mm}$. On prélève des échantillons de taille n avec $n \geqslant 30$. Soit $\overline{X}$ la variable qui à chaque échantillon de n pièces associe la moyenne de longueurs.

1. Quelle est la loi suivie par $\overline{X}$? Précisez les paramètres de cette loi.
2. Trouvez la taille n de l'échantillon pour que $p\left(4,17 < \overline{X} < 4,23\right) = 0,95$.
3. Trouvez la taille n de l'échantillon pour que $p\left(4,17 < \overline{X} < 4,23\right) = 0,99$.

II. 2) Loi des fréquences

Dans une population la proportion p d'individus possèdent une propriété.
On prélève aléatoirement un échantillon de taille n.
F est la fréquence de la propriété dans l'échantillon $\Rightarrow$ F suit **approximativement** la loi $\mathcal{N}$ de moyenne p et d'écart-type $\sqrt{\dfrac{p(1-p)}{n}}$

II. 2) Loi des fréquences

Dans une population la proportion p d'individus possèdent une propriété. On prélève aléatoirement un échantillon de taille n.
F est la fréquence de la propriété dans l'échantillon $\Rightarrow$ F suit **approximativement** la loi $\mathcal{N}$ de moyenne p et d'écart-type $\sqrt{\dfrac{p(1-p)}{n}}$

Exemple : Lors d'un référendum, 53 % des électeurs vont voter OUI. On prélève un échantillon de 1 000 électeurs. Quelle est la probabilité qu'il y ait entre 52,5 % et 53,5 % d'électeurs votant OUI dans cet échantillon ?

II. 2) Loi des fréquences

Dans une population la proportion p d'individus possèdent une propriété.
On prélève aléatoirement un échantillon de taille n.
F est la fréquence de la propriété dans l'échantillon $\Rightarrow$ F suit **approximativement** la loi $\mathcal{N}$ de moyenne p et d'écart-type $\sqrt{\dfrac{p(1-p)}{n}}$

Exemple : Lors d'un référendum, 53 % des électeurs vont voter OUI. On prélève un échantillon de 1 000 électeurs. Quelle est la probabilité qu'il y ait entre 52,5 % et 53,5 % d'électeurs votant OUI dans cet échantillon ?

Soit F la variable aléatoire représentant le pourcentage d'électeurs de l'échantillon votant OUI. F suit la loi $\mathcal{N}$ de moyenne $0,53$ et d'écart-type $\sqrt{\frac{0{,}53\times 0{,}46}{1000}} \simeq 0,0158$. La probabilité recherchée est donc :

$$p(0,525 \leqslant F \leqslant 0,535) \simeq 0,2472$$

Exercice 4

Une entreprise produit une grande série de pièces pour le bâtiment. Pour analyser la qualité de la production, on prélève au hasard un échantillon de $n = 64$ pièces. On suppose que le pourcentage de pièces défectueuses dans la production est de 4 %. Soit F la variable aléatoire qui à tout échantillon de n pièces associe la fréquence de pièces défectueuses dans cet échantillon.

1. quelle est, approximativement, la loi suivie par F ? Précisez-en les paramètres.
2. Quel est la probabilité que sur un échantillon donné, la fréquence de pièce défectueuses soit comprise entre 3 % et 5 % ?
3. Déterminez le réel a tel que $p(0,04 - a \leqslant F \leqslant 0,04 + a) = 0,95$.
4. Quelle valeur doit-on donner à n pour que $p(0,03 \leqslant F \leqslant 0,05) = 0,99$

Exercice 5

Dans une population, on constate qu'il y a 28 % de fumeurs. On prend un échantillon de 400 personnes.

1. Quel est la probabilité d'avoir dans cet échantillon un pourcentage de fumeurs compris entre 26 % et 30 % de fumeurs ?
2. Quelle est la probabilité d'avoir dans cet échantillon un pourcentage de fumeurs inférieur à 22 % ?

III. Estimation

On a réalisé un échantillon. On connaît les paramètres de l'échantillon et on veut estimer ceux de la population.

III. 1) Estimation ponctuelle

Dans un échantillon de taille n on a relevé la moyenne $\overline{x_e}$ et l'écart-type σ_e.

ou bien on a relevé la fréquence f_e.

III. 1) Estimation ponctuelle

Dans un échantillon de taille n on a relevé la moyenne $\overline{x_e}$ et l'écart-type σ_e.

ou bien on a relevé la fréquence f_e.

1. La moyenne de la population est estimée par $\overline{x_e}$.
2. La fréquence dans la population est estimée par f_e.
3. L'écart-type σ dans la population est estimée par $s = \sqrt{\frac{n}{n-1}}\sigma_e$.

Le dernier montre que la mesure de σ_e sous-estime le σ réel : Il y a en moyenne moins de diversité dans un échantillon que dans la population entière !

Exemple

À un examen auquel se présentent 100 000 candidats, 100 copies ont été corrigées. La moyenne sur ce paquet était de 12 avec un écart-type de 3. Dans ce paquet, 5 copies étaient blanches.

Exemple

À un examen auquel se présentent 100 000 candidats, 100 copies ont été corrigées. La moyenne sur ce paquet était de 12 avec un écart-type de 3. Dans ce paquet, 5 copies étaient blanches.

On peut estimer que sur l'ensemble des copies :

1. la moyenne était de 12,
2. l'écart-type $\sigma = \sqrt{\frac{100}{99}}3 \simeq 3,015$,
3. 5% de l'ensemble des copies étaient blanches.

III. 2) Estimation d'une moyenne par intervalle de confiance

On se donne un taux de confiance $1-\alpha$ (risque de α). On peut déduire t tel que $2\Pi(t)-1=1-\alpha$ (voir première partie)

Au taux de confiance $1-\alpha$, l'intervalle de confiance de la moyenne est :

$$I_C = \left[\overline{x_e} - t\frac{\sigma}{\sqrt{n}} ; \overline{x_e} + t\frac{\sigma}{\sqrt{n}}\right]$$

$\overline{x_e}$ est la moyenne de l'échantillon considéré, σ est l'écart-type de la population (s'il est connu) où son estimation ponctuelle s.

III. 2) Estimation d'une moyenne par intervalle de confiance

On se donne un taux de confiance $1-\alpha$ (risque de α). On peut déduire t tel que $2\Pi(t)-1=1-\alpha$ (voir première partie)

Au taux de confiance $1-\alpha$, l'intervalle de confiance de la moyenne est :

$$I_C = \left[\overline{x_e} - t\frac{\sigma}{\sqrt{n}} ; \overline{x_e} + t\frac{\sigma}{\sqrt{n}}\right]$$

$\overline{x_e}$ est la moyenne de l'échantillon considéré, σ est l'écart-type de la population (s'il est connu) où son estimation ponctuelle s.

Dans l'exemple précédent, $\overline{x_e}=12$ et $\sigma \simeq 3,015$. Si on prend le taux de confiance 95 %, on aura $t=1,96$ et

$$I_C \simeq \left[12 - 1,96\frac{3,015}{\sqrt{100}} ; 12 + 1,96\frac{3,015}{\sqrt{100}}\right] \simeq [11,409 ; 12,591]$$

Exercice 6

Une entreprise fabrique des tiges métalliques. Leur longueur et leur diamètre sont exprimés en millimètres. On prélève au hasard un échantillon de 50 tiges dans la production d'une journée. Soit $\overline{D}$ la variable aléatoire qui, à tout échantillon de 50 tige, associe le diamètre moyen de ces tiges. On suppose que $\overline{D}$ suit la loi $\mathcal{N}$ de moyenne *mu* et écart-type $\frac{\sigma}{\sqrt{50}}$, avec $\sigma = 0,19$ et μ inconnu. Sur l'échantillon, on a observé $\mu_e = 9,99$.

1. À partir de ces informations, donnez une estimation ponctuelle de la moyenne μ des diamètres des tiges produites dans une journée.
2. Déterminez un intervalle de confiance de la moyenne μ à 95%.
3. On considère l'affirmation : « μ appartient obligatoirement à l'intervalle de la question précédente ». Cette affirmation est-elle vraie ?

Exercice 7

On a contrôlé le dosage d'un produit dans un mélange à la sortie d'une chaîne de conditionnement. On a prélevé un échantillon de 100 lots de 5 kg dans la production d'une journée. On a obtenu les résultats suivants, où P_i représente la masse de produit en grammes, et n_i l'effectif correspondant.

P_i	142	144	146	148	150	152	154	156	158	160
n_i	1	5	6	21	32	22	7	4	1	1

1. Calculez μ_e et σ_e.
2. Déduisez-en une estimation ponctuelle de μ et σ de masse de produit dans la production de la journée.
3. Donnez un intervalle de confiance de la moyenne μ avec le coefficient de confiance de 5%.
4. même question avec un coefficient de confiance de 99 %.

III. 3) Estimation d'une fréquence par intervalle de confiance

Au taux de confiance $1 - \alpha$, l'intervalle de confiance de la fréquence est :

$$I_C = \left[f - t\sqrt{\frac{f(1-f)}{n-1}} ; f + t\sqrt{\frac{f(1-f)}{n-1}} \right]$$

f est la fréquence du caractère dans l'échantillon considéré.
On note généralement p la fréquence dans l'ensemble de la population.

III. 3) Estimation d'une fréquence par intervalle de confiance

Au taux de confiance $1-\alpha$, l'intervalle de confiance de la fréquence est :

$$I_C = \left[f - t\sqrt{\frac{f(1-f)}{n-1}}\,;\, f + t\sqrt{\frac{f(1-f)}{n-1}} \right]$$

f est la fréquence du caractère dans l'échantillon considéré.
On note généralement p la fréquence dans l'ensemble de la population.

Dans l'exemple du paquet de copies : $n = 100$ copies, $f = 5\,\%$ de copies blanches. On souhaite déterminer un intervalle de confiance de p au seuil de confiance de 95 %.

$$I_C = \left[0,05 - 1,96\sqrt{\frac{0,05\times 0,95}{99}}\,;\, 0,05 + 1,96\sqrt{\frac{0,05\times 0,95}{99}}\right]$$

$$I_C \simeq [0,007\,;\,0,093]$$

Exercice 8

Dans une ville, la proportion p de ménages qui possèdent au moins un téléviseur est inconnue.

On interroge 400 ménages. 327 de ces ménages possèdent au moins un téléviseur.

1. Donnez une estimation ponctuelle de p.
2. Donnez une estimation de p par intervalle de confiance avec le risque 5%.
3. Quelle taille d'échantillon aurait-il fallu prendre pour obtenir un intervalle de confiance à $\pm 2\,\%$, avec un taux de confiance de $99\,\%$?

IV. Test de validité

IV. 1) Principe général

Il s'agit de prendre une décision sur l'ensemble de la population à partir d'une étude sur des échantillons. On émet une hypothèse qui doit être confirmée ou rejetée.

IV. 1) Principe général

Il s'agit de prendre une décision sur l'ensemble de la population à partir d'une étude sur des échantillons. On émet une hypothèse qui doit être confirmée ou rejetée.

Méthode : Une variable aléatoire étant définie, **avant** le prélèvement d'un ou plusieurs échantillons :

Il s'agit de prendre une décision sur l'ensemble de la population à partir d'une étude sur des échantillons. On émet une hypothèse qui doit être confirmée ou rejetée.

Méthode : Une variable aléatoire étant définie, **avant** le prélèvement d'un ou plusieurs échantillons :

1. On choisit l'hypothèse nulle H_0 et l'hypothèse alternative H_1.

IV. 1) Principe général

Il s'agit de prendre une décision sur l'ensemble de la population à partir d'une étude sur des échantillons. On émet une hypothèse qui doit être confirmée ou rejetée.

Méthode : Une variable aléatoire étant définie, **avant** le prélèvement d'un ou plusieurs échantillons :

1. On choisit l'hypothèse nulle H_0 et l'hypothèse alternative H_1.
2. On détermine la région critique à un seuil α donné.

IV. 1) Principe général

Il s'agit de prendre une décision sur l'ensemble de la population à partir d'une étude sur des échantillons. On émet une hypothèse qui doit être confirmée ou rejetée.

Méthode : Une variable aléatoire étant définie, **avant** le prélèvement d'un ou plusieurs échantillons :

1. On choisit l'hypothèse nulle H_0 et l'hypothèse alternative H_1.
2. On détermine la région critique à un seuil α donné.
3. On énonce la règle de décision :
   - Si un paramètre de l'échantillon est dans la région critique, on rejette H_0 ($\Rightarrow$ on accepte H_1),
   - sinon on accepte H_0.

IV. 2) Test bilatéral

On teste l'affirmation $\overline{x} = \mu_{ESTIM}$

Pour simplifier, dans ce cas, on suppose σ connu.

Ce pourrait être $p = p_{ESTIM}$ pour une fréquence.

IV. 2) Test bilatéral

On teste l'affirmation $\overline{x} = \mu_{ESTIM}$

Pour simplifier, dans ce cas, on suppose σ connu.

Ce pourrait être $p = p_{ESTIM}$ pour une fréquence.

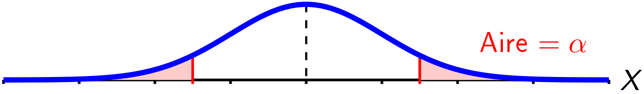
1. $H_0 : \overline{x} = \mu_{ESTIM}$ et $H_1 : \overline{x} \neq \mu_{ESTIM}$

On teste l'affirmation $\overline{x} = \mu_{ESTIM}$

Pour simplifier, dans ce cas, on suppose σ connu.

Ce pourrait être $p = p_{ESTIM}$ pour une fréquence.

1. H_0 : $\overline{x} = \mu_{ESTIM}$ et H_1 : $\overline{x} \neq \mu_{ESTIM}$
2. Zone critique (ZC) au seuil α :

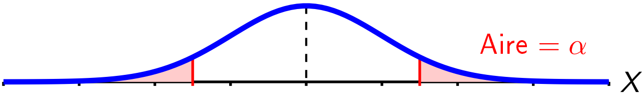


On teste l'affirmation $\overline{x} = \mu_{ESTIM}$

Pour simplifier, dans ce cas, on suppose σ connu.

Ce pourrait être $p = p_{ESTIM}$ pour une fréquence.

1. $H_0 : \overline{x} = \mu_{ESTIM}$ et $H_1 : \overline{x} \neq \mu_{ESTIM}$
2. Zone critique (ZC) au seuil α :



3. On énonce le test :
   - $X \in ZC \Rightarrow$ rejet de H_0
   - $X \notin ZC \Rightarrow$ acceptation de H_0

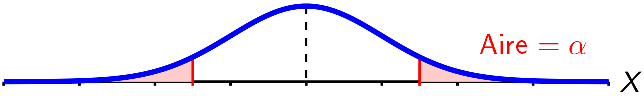
IV. 2) Test bilatéral

On teste l'affirmation $\overline{x} = \mu_{ESTIM}$

Pour simplifier, dans ce cas, on suppose σ connu.

Ce pourrait être $p = p_{ESTIM}$ pour une fréquence.

1. $H_0 \ : \overline{x} = \mu_{ESTIM}$ et $H_1 \ : \overline{x} \neq \mu_{ESTIM}$
2. Zone critique (ZC) au seuil α :



3. On énonce le test :
   - $X \in ZC \Rightarrow$ rejet de H_0
   - $X \notin ZC \Rightarrow$ acceptation de H_0
4. On réalise le test.

Exercice 9

Une usine fabrique des roues dentées dont le diamètre annoncé est de 23,65 mm.

Un client commande un lot et veut vérifier que les roues dentées ont bien le diamètre annoncé. Il prend un lot de 100 roues et obtient une moyenne $\overline{x_e} = 23{,}644\,\mathrm{mm}$ et un écart-type $\sigma_e = 0{,}018\,\mathrm{mm}$.

Le client estime que l'écart-type de l'échantillon est une bonne estimation de la totalité du lot. Peut-il admettre, au risque de 5 %, l'affirmation du fournisseur ?

IV. 3) Test unilatéral

Dans le test précédent, on envisageait deux risques : que $\overline{x}$ soit sensiblement $>$ ou $<$ à la valeur μ_{ESTIM}.

Maintenant, on ne va envisager qu'un seul des deux risques.

IV. 3) Test unilatéral

Dans le test précédent, on envisageait deux risques : que $\overline{x}$ soit sensiblement $>$ ou $<$ à la valeur μ_{ESTIM}.

Maintenant, on ne va envisager qu'un seul des deux risques.

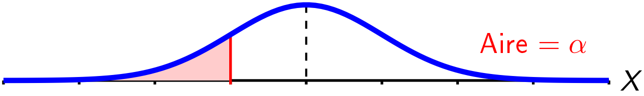
1. $H_0 \ : \overline{x} = \mu_{ESTIM}$ et $H_1 \ : \overline{x} < \mu_{ESTIM}$

IV. 3) Test unilatéral

Dans le test précédent, on envisageait deux risques : que $\overline{x}$ soit sensiblement $>$ ou $<$ à la valeur μ_{ESTIM}.

Maintenant, on ne va envisager qu'un seul des deux risques.

1. $H_0 : \overline{x} = \mu_{ESTIM}$ et $H_1 : \overline{x} < \mu_{ESTIM}$
2. Zone critique (ZC) au seuil α :

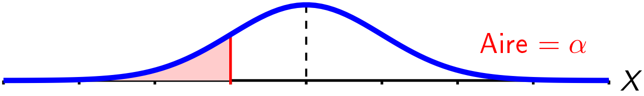


IV. 3) Test unilatéral

Dans le test précédent, on envisageait deux risques : que $\overline{x}$ soit sensiblement $>$ ou $<$ à la valeur μ_{ESTIM}.

Maintenant, on ne va envisager qu'un seul des deux risques.

1. $H_0 : \overline{x} = \mu_{ESTIM}$ et $H_1 : \overline{x} < \mu_{ESTIM}$
2. Zone critique (ZC) au seuil α :

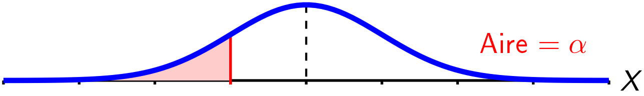


3. On énonce le test :
   - $X \in$ ZC $\Rightarrow$ rejet de $H_0 \Rightarrow$ Acceptation de H_1
   - $X \notin$ ZC $\Rightarrow$ acceptation de H_0

Dans le test précédent, on envisageait deux risques : que $\overline{x}$ soit sensiblement $>$ ou $<$ à la valeur μ_{ESTIM}.

Maintenant, on ne va envisager qu'un seul des deux risques.

1. $H_0 \ : \overline{x} = \mu_{ESTIM}$ et $H_1 \ : \overline{x} < \mu_{ESTIM}$
2. Zone critique (ZC) au seuil α :



3. On énonce le test :
   - $X \in$ ZC $\Rightarrow$ rejet de H_0 $\Rightarrow$ Acceptation de H_1
   - $X \notin$ ZC $\Rightarrow$ acceptation de H_0
4. On réalise le test.

On aurait pu envisager le risque $H_1 \ : \overline{x} > \mu_{ESTIM}$ et donc prendre une ZC symétrique.

Exercice 10

Une entreprise a constaté qu'un certain nombre de ballons d'eau chaude qu'elle installe ont un défaut de fonctionnement. On s'intéresse à la fréquence de ballons installés ayant un défaut.

L'entreprise installe 304 ballons et constate que 18 ont un défaut.

Peut-on affirmer, au seuil de risque de 1 % que p, la fréquence de défaut, est supérieur à 5 % ?

IV. 4) Risque de première espèce

Supposons que suite à un test, on rejette H_0 alors que H_0 est vraie. C'est le risque de première espèce, risque α.

IV. 4) Risque de première espèce

Supposons que suite à un test, on rejette H_0 alors que H_0 est vraie. C'est le risque de première espèce, risque α.

Exemple : p vaut vraiment 60 % mais je ne le sais pas. Je décide de prélever un échantillon de $n = 200$ et je décide d'un test au risque $\alpha = 5\,\%$:

$$I_F \simeq [53,2\,\%\,;66,8\,\%]$$

IV. 4) Risque de première espèce

Supposons que suite à un test, on rejette H_0 alors que H_0 est vraie. C'est le risque de première espèce, risque α.

Exemple : p vaut vraiment 60 % mais je ne le sais pas. Je décide de prélever un échantillon de $n = 200$ et je décide d'un test au risque $\alpha = 5\,\%$:

$$I_F \simeq [53,2\,\%\,;66,8\,\%]$$

Se peut-il que dans un échantillon, $f \notin I_F$, me laissant croire que $p \neq 60\,\%$ alors que $p = 60\,\%$ en réalité ?

IV. 4) Risque de première espèce

Supposons que suite à un test, on rejette H_0 alors que H_0 est vraie. C'est le risque de première espèce, risque α.

Exemple : p vaut vraiment 60 % mais je ne le sais pas. Je décide de prélever un échantillon de $n = 200$ et je décide d'un test au risque $\alpha = 5\,\%$:

$$I_F \simeq [53,2\,\%\,;66,8\,\%]$$

Se peut-il que dans un échantillon, $f \notin I_F$, me laissant croire que $p \neq 60\,\%$ alors que $p = 60\,\%$ en réalité ?

OUI ! Si $p = 60\,\%$, la probabilité que $f \notin I_F$ est précisément $\alpha = 5\,\%$, c'est pour cela que l'on dit « Intervalle de fluctuation au risque α »

IV. 5) Risque de seconde espèce

Supposons que suite à un test, on accepte H_0 alors que H_0 est fausse. C'est le risque de seconde espèce, risque β.

Supposons que suite à un test, on accepte H_0 alors que H_0 est fausse. C'est le risque de seconde espèce, risque β.

Exemple : p ne vaut pas 60 % mais je fais l'hypothèse que c'est le cas. Je décide de prélever un échantillon de $n = 200$ et je décide d'un test au risque $\alpha = 5\,\%$:

$$I_F \simeq [53,2\,\%\,;66,8\,\%]$$

IV. 5) Risque de seconde espèce

Supposons que suite à un test, on accepte H_0 alors que H_0 est fausse. C'est le risque de seconde espèce, risque β.

Exemple : p ne vaut pas 60 % mais je fais l'hypothèse que c'est le cas. Je décide de prélever un échantillon de $n = 200$ et je décide d'un test au risque $\alpha = 5\,\%$:

$$I_F \simeq [53,2\,\%\,;66,8\,\%]$$

Se peut-il que dans un échantillon, $f \in I_F$, me laissant croire que $p = 60\,\%$ alors que $p \neq 60\,\%$ en réalité ?

IV. 5) Risque de seconde espèce

Supposons que suite à un test, on accepte H_0 alors que H_0 est fausse. C'est le risque de seconde espèce, risque β.

Exemple : p ne vaut pas 60 % mais je fais l'hypothèse que c'est le cas. Je décide de prélever un échantillon de $n = 200$ et je décide d'un test au risque $\alpha = 5\,\%$:

$$I_F \simeq [53,2\,\% \,; 66,8\,\%]$$

Se peut-il que dans un échantillon, $f \in I_F$, me laissant croire que $p = 60\,\%$ alors que $p \neq 60\,\%$ en réalité ?

OUI ! Mais le calcul est plus délicat.
Il faut envisager, pour chaque valeur de $p \neq 60\,\%$ possible, la probabilité que $f \in I_F$.