TD 3 - Exploitation et traitement des données

<u>Exercice 1</u> (Rappels : moyenne, écart-type)

a) Evaluer la moyenne arithmétique de la série de nombres par deux méthodes :

5, 3, 6, 5, 4, 5, 2, 8, 6, 5, 4, 8, 3, 4, 5, 4, 8, 2, 5, 8.

$\bar{x} = 5.0$.

Donner l'expression analytique correspondante à chacune des deux méthodes.

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

$\bar{x} = \frac{1}{n}\sum_{k=1}^{K} n_k\, x^k$ et $\sum_{k=1}^{K} n_k = n$.

Avec :

k	**1**	**2**	**3**	**4**	**5**	**6**		
x^k	2	3	4	5	6	8		
n_k	2	2	4	6	2	4		$\sum_{k=1}^{K} n_k$=20

b) Calculer pour les données du a) la variance non corrigée en 1/n et l'écart-type correspondant, après avoir écrit la formule analytique pour deux méthodes.

$s_n = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sqrt{3.40} = 1.84.$

c) Rappeler la formule pour la variance corrigée (sans biais) en 1/(n-1) et l'écart-type correspondant. Les calculer pour les données ci-dessus.

$s_n^* = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sqrt{3.58} = 1.89.$

d) Un échantillon de 4 mesures de diamètres d'une sphère comprend les résultats suivants : 6.33, 6.37, 6.36, 6.32 et 6.37. Calculer les estimations non biaisées (et efficaces) de la vraie moyenne (celle de la population) et de la vraie variance (celle de la population).

$\bar{x} = 6.35$ et $s_n^* = \sqrt{0.00055} = 0.02.$

<u>Exercice 2</u> (Intervalles de confiance, loi de Gauss)

Indication : Il est à noter que pour n assez grand supérieur à 30, on suppose que les expressions avec la loi (normale) de Gauss deviennent correctes, donc le vrai écart-type σ peut être remplacé par son estimation corrigée (dit également modifié) s_n (non biaisée) en 1/(n-1), il est inutile de préférer une loi de Student, voir le formulaire ci-joint pour *n large*.

Les mesures des diamètres d'un échantillon aléatoire de 200 billes de roulement fabriquées par une machine donnée au cours d'une semaine présentent une moyenne $\bar{x}$ de 0.824 cm et d'écart-type corrigé s_n de 0.042 cm pour cet échantillon.

- Evaluer un intervalle de confiance à 95% puis à 99% pour la vraie moyenne, après avoir récupéré (en annexe) les valeurs z_α correspondantes pour les risques a) α=5% et b) α=1%.

  D'après le cours, l'intervalle de confiance s'écrit : $I_\alpha = \bar{x} \mp z_\alpha \frac{\sigma}{\sqrt{n}}$ en approximant le vrai écart-type σ par s_n car n est large (200>30).

  On sait donc que il y a le résultat de probabilité suivant : P(μϵI_α)=1-α
  Typiquement, α=0.05 donc 1-α=95% pour la probabilité de trouver la vraie moyenne inconnue dans l'intervalle de confiance I_α.

  D'où le calcul en remplaçant les valeurs dans l'expression de l'intervalle:

  - Pour α=5%, on lit en annexe la valeur $z_\alpha = 1.96$ donc : $I_{0.05} = 0.824 \mp 0.0058$ cm.

  - Pour α=1%, on lit en annexe la valeur $z_\alpha = 2.58$ donc : $I_{0.01} = 0.824 \mp 0.0077$ cm.

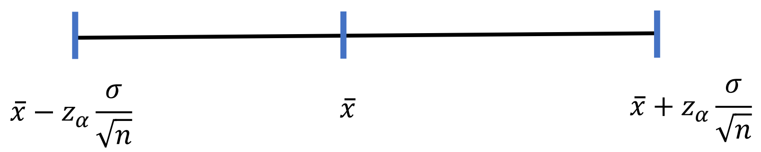
  Pour rappel (voir cours), l'expression obtenue pour l'intervalle provient de la standardisation de la variable aléatoire moyenne, si bien que la statistique considérée suit une loi normale N(0,1):

  $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$.
  Donc les quantiles z_α sont obtenus à l'aide de la loi centrée-réduite N(0,1), si bien que :

  $P(-z_\alpha \leq Z \leq z_\alpha) = 1 - \alpha$.

- Dessiner graphiquement l'intervalle de confiance.

Par exemple,



Exercice 3 (Intervalles de confiance, loi de Student)

Un échantillon constitué par 10 mesures de diamètre d'une sphère présente une moyenne $\bar{x}$ de 4.38 cm et un écart-type non corrigé de 0.06 cm.

Calculer l'intervalle de confiance pour la valeur réelle µ avec le risque de a) 5% puis b) 1%, après avoir récupéré (en annexe) les valeurs z_α correspondantes.

Idem que précédemment, sauf que la statistique est maintenant, $T = \frac{\bar{X}-\mu}{S_n/\sqrt{n}} \sim \mathcal{T}(n-1)$.

Donc les quantiles changent, et différent du cas normal, et même dépendent de n (sauf lorsque n devient large car la loi de Student devient alors une loi normale centrée réduite).

Au lieu de z_α on a donc $t_{n-1;\alpha}$ qui dépend de n tout comme la loi $\mathcal{T}(n-1)$ donc :

$P(-t_{n-1;\alpha} \leq T \leq t_{n-1;\alpha}) = 1 - \alpha.$

Pour n=10, on lit en annexe les valeurs $t_{9;0.05} = 2.262$ et $t_{9;0.01} = 3.250$.

D'où l'intervalle, par exemple

- Pour α=5%, $I_\alpha = \bar{x} \mp t_{n-1;\alpha} \frac{s_n^*}{\sqrt{n}} = \bar{x} \mp t_{n-1;\alpha} \frac{s_n}{\sqrt{n-1}} = [4.33; 4.43]$cm
- Pour α=1%, $I_\alpha = \bar{x} \mp t_{n-1;\alpha} \frac{s_n^*}{\sqrt{n}} = \bar{x} \mp t_{n-1;\alpha} \frac{s_n}{\sqrt{n-1}} = [4.31; 4.45]$cm

Exercice 4 (intervalle de confiance, proportion)

On a obtenu 24 coté « face » au cours de 40 lancers d'une pièce.

a) Quelle est la valeur estimée $\hat{p}$ de la vraie proportion inconnue p d'après l'échantillon ?

b) Calculer un intervalle de confiance pour p à 95% pour la vraie fréquence (ou proportion) p de cotés « face » obtenues au cours d'un nombre illimité de lancers.

Idem que précédemment pour le cas normal, sauf que la statistique est maintenant asymptotique pour n suffisamment large (typiquement n>40, np>5, n(1-p)>5) :

$Z = \frac{F-p}{\sqrt{p(1-p)}/\sqrt{n}} \sim N(0,1).$

Avec ici F, la fréquence aléatoire qui en réalité s'identifie à la moyenne précédente $\bar{X}_n$.

Le même raisonnement aboutit donc là encore, avec $\hat{p} = \frac{24}{40} = 0.6$,à l'intervalle (voir cours) :

$I_\alpha = \hat{p} \mp z_\alpha \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = [0.44; 0.76]$.

Exercice 6 (Test d'hypothèse, deux moyennes)

Deux groupes de 40 et 60 étudiants respectivement, ont subi le même examen. La note moyenne du premier groupe a été de 74 avec un écart-type de 8, et pour le second, de 78 avec un écart-type de 7. Existe-t-il une différence significative entre les résultats moyens des deux groupes au risque a) 5% et b) 1% ?

On veut tester :

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

La statistique en jeu est donc, avec l'hypothèse $\mu_1 = \mu_2$:

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1}{n_1} + \frac{\sigma_2}{n_2}}} \sim N(0,1).$$

D'où le calcul en approchant les vrais écart-types par les estimations:

$$z = \frac{\bar{x}_1 - x_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = -2.57 .$$

Comme z=-2.57 n'appartient pas à l'intervalle [-1.96 ;1.96], on ne peut accepter l'hypothèse H_0 donc on en déduit que la différence des moyennes est significative au risque 5%. Ensuite, l'intervalle devient [-2.58 ;2.58] pour le risque 1%, donc l'hypothèse H_0 ne peut être rejetée puisque z est maintenant dans la région d'acceptation.

Comme la valeur z=-2.57 est au bord de l'intervalle de la région d'acceptation à 1%, la décision d'acceptation demande une investigation si possible du résultat. Pour certains statisticiens, des décisions au niveau 1% sont hautement significatifs, des résultats significatifs au niveau 5% mais pas au niveau 1% ne sont que probablement significatifs et enfin des résultats significatifs à des niveaux supérieur à 5% sont pas significatifs.

Exercice 7 (Test d'hypothèse, une moyenne)

Dans le passé, une machine a produit des pièces d'un épaisseur moyenne de 0.050 cm. Pour déterminer si cette machine fonctionne toujours correctement, on choisit un échantillon de 10 pièces or l'épaisseur moyenne est de 0.053 cm pour cette série tandis que l'écart-type est de 0.003 cm. Peut-on soutenir son affirmation à un niveau de risque a) 0.05 ou b) 0.01 .

Le test s'écrit :

$H_0 : \mu = 0.05$

$H_1 : \mu \neq 0.05$

La statistique s'écrit :

$T = \frac{\bar{X}-\mu}{S_n/\sqrt{n}} \sim \mathcal{T}(n-1)$ car n est petit.

Remarquons que la distribution des observations est supposée normale.

Le calcul sur les données donne donc, avec l'hypothèse H_0 (en corrigeant l'écart-type):

$t = \frac{0.053-0.05}{0.003/\sqrt{10-1}}$ =3.

Il suffit maintenant de vérifier avec les intervalles probables de trouver t en fonction des risques α, on a le même quantile que pour l'exercice 3, $t_{9;0.05} = 2.262$ et $t_{9;0.01} = 3.250$ d'où :

- Cas 5%, t ∉ [-2.26 ;2.26] donc l'hypothèse nulle H_0 n'est pas acceptée.
- Cas 1%, t ∈ [-3.25 ;3.25] donc l'hypothèse nulle H_0 est acceptée.

En résumé, l'hypothèse est *probablement significative*, mais des vérifications s'imposent.

<u>Exercice 8</u> (Intervalle de confiance pour une différence)

Un échantillon de 150 ampoules de marque A présente une durée de vie moyenne de 1400 heures avec un écart-type de 120 heures. Un échantillon de 200 ampoules de marque B présente une durée de vie moyenne de 1200 heures avec un écart-type de 80 heures. Evaluer un intervalle de confiance à 99% pour les différences de durée de vie moyenne des deux populations.

Pour résoudre l'exercice, on sait que l'écart-type de $D = \bar{X}_1 - \bar{X}_2$ est $\sigma_D = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ en supposant $X_1 \sim N(\mu_1, \sigma_1\ )$ et $X_2 \sim N(\mu_2, \sigma_1\ )$ car les variables aléatoires sont toutes indépendantes :

$E(D) = E[\bar{X}_1 - \bar{X}_2] = \mu_1 - \mu_2$.

$V(D) = V[\bar{X}_1 - \bar{X}_2] = V[\bar{X}_1] + V[\bar{X}_2] = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$.

Donc en standardisant une loi normale d'espérance $\mu_1 - \mu_2$ et d'écart-type $\sigma_D = \sqrt{V(D)}$, on retrouve une expression similaire à celle pour l'intervalle d'une moyenne d'un seul échantillon. En effet, soit :

$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sigma_D} \sim N(0,1)$.

Donc,

$P\left(-2.58 \leq \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sigma_D} \leq 2.58\right) = 0.99$.

On a la probabilité de 99% pour la variable aléatoire D=$\bar{X}_1 - \bar{X}_2$ d'avoir le résultat suivant :

$-2.58 \leq \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sigma_D} \leq 2.58$.

$(\bar{X}_1 - \bar{X}_2) - 2.58\,\sigma_D \leq \mu_1 - \mu_2 \leq (\bar{X}_1 - \bar{X}_2) + 2.58\,\sigma_D$.

Donc $\mu_1 - \mu_2$ est dans l'intervalle suivant à 99%, en passant aux valeurs d'échantillon :

$\bar{x}_1 - \bar{x}_2 \mp z_\alpha\,\sigma_D = (1400\text{-}1200) \mp z_\alpha\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 200 \mp 2.58 * 11.31 = [170.80; 229.19]$.

Exercice 9 (Test pour une moyenne)

Dans une expérience sur l'acuité visuelle, on a demandé à 14 individus d'évaluer la distance d'un objet placé à 20 cm. Il est obtenu les résultats, 17, 20, 21, 14, 18, 19, 19, 24, 21, 16, 23, 15, 21, et 20. Peut-on affirmer que les individus ont en moyenne de la difficulté à évaluer correctement la vraie distance ?

Les hypothèses du test s'écrivent,

H_0 : μ=20

H_1 : μ≠20

En suivant la résolution de l'exercice 7, on écrit avec la variance corrigée:

t = $\frac{19.14-20}{2.91/\sqrt{14}}$ = -1.11 .

Comme l'intervalle probable à 95% est égal à [-2.16 ; 2.16] pour n=14, on ne peut rejeter H_0.

Exercice 10 (Test pour une proportion)

Comment tester si une pièce avec pile ou face est bien équilibrée après 100 lancers ?

Il faut tester si l'hypothèse H_0 : p=0.5 est acceptable ou non, lorsque confrontée à H_1 : p≠0.5. Ensuite, compter le nombre de piles ou faces obtenu, en déduire une proportion $\hat{p}$ puis calculer la statistique pour la décision. Il s'agit d'une observation de $Z = \frac{F-p}{\sqrt{p(1-p)}/\sqrt{n}} \sim N(0,1)$, donc on remplace F par la proportion estimée $\hat{p}$ et la vraie proportion p par 0.5 puis vérifie si la statistique résultante z tombe dans la zone d'acceptation ou la zone de rejet.

ANNEXE

Paramater estimation

❒ **Random sample** – A random sample is a collection of n random variables $X_1, ..., X_n$ that are independent and identically distributed with X.

❒ **Estimator** – An estimator $\hat{\theta}$ is a function of the data that is used to infer the value of an unknown parameter θ in a statistical model.

❒ **Bias** – The bias of an estimator $\hat{\theta}$ is defined as being the difference between the expected value of the distribution of $\hat{\theta}$ and the true value, i.e.:

$$\boxed{\mathrm{Bias}(\hat{\theta}) = E[\hat{\theta}] - \theta}$$

Remark: an estimator is said to be unbiased when we have $E[\hat{\theta}] = \theta$.

❒ **Sample mean and variance** – The sample mean and the sample variance of a random sample are used to estimate the true mean μ and the true variance σ^2 of a distribution, are noted $\overline{X}$ and s^2 respectively, and are such that:

$$\boxed{\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i} \quad \text{and} \quad \boxed{s^2 = \hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2}$$

❒ **Central Limit Theorem** – Let us have a random sample $X_1, ..., X_n$ following a given distribution with mean μ and variance σ^2, then we have:

$$\boxed{\overline{X} \underset{n \to +\infty}{\sim} \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)}$$

Confidence intervals

❒ **Confidence level** – A confidence interval $CI_{1-\alpha}$ with confidence level $1-\alpha$ of a true parameter θ is such that $1-\alpha$ of the time, the true value is contained in the confidence interval:

$$\boxed{P(\theta \in CI_{1-\alpha}) = 1 - \alpha}$$

❒ **Confidence interval for the mean** – When determining a confidence interval for the mean μ, different test statistics have to be computed depending on which case we are in. The following table sums it up:

Distribution	Sample size	σ^2	Statistic	$1-\alpha$ confidence interval
$X_i \sim \mathcal{N}(\mu, \sigma)$	any	known	$\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0,1)$	$\left[\overline{X} - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$
	small	unknown	$\frac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}} \sim t_{n-1}$	$\left[\overline{X} - t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}, \overline{X} + t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}\right]$
$X_i \sim$ any	large	known	$\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0,1)$	$\left[\overline{X} - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$
		unknown	$\frac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}} \sim \mathcal{N}(0,1)$	$\left[\overline{X} - z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}, \overline{X} + z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}\right]$
$X_i \sim$ any	small	any	Go home!	Go home!

Hypothesis testing

❒ **Errors** – In a hypothesis test, we note α and β the type I and type II errors respectively. By noting T the test statistic and R the rejection region, we have:

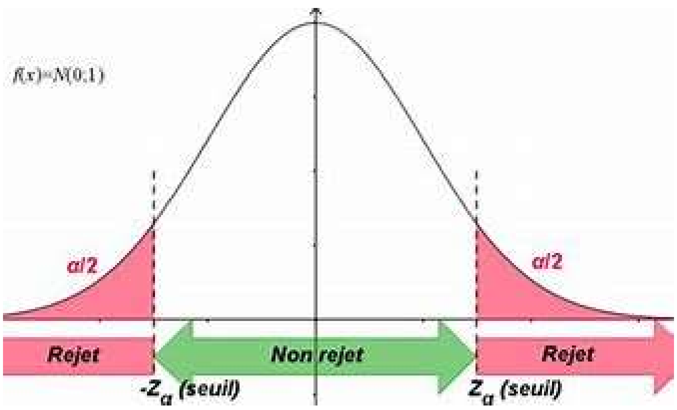
$$\boxed{\alpha = P(T \in R | H_0 \text{ true})} \quad \text{and} \quad \boxed{\beta = P(T \notin R | H_1 \text{ true})}$$

❒ **p-value** – In a hypothesis test, the p-value is the probability under the null hypothesis of having a test statistic T at least as extreme as the one that we observed T_0. We have:

Case	Left-sided	Right-sided	Two-sided
p-value	$P(T \leqslant T_0 \| H_0 \text{ true})$	$P(T \geqslant T_0 \| H_0 \text{ true})$	$P(\|T\| \geqslant \|T_0\| \| H_0 \text{ true})$

Distribution of X_i, Y_i	n_X, n_Y	σ_X^2, σ_Y^2	Statistic
Normal	any	known	$\dfrac{(\overline{X}-\overline{Y})-\delta}{\sqrt{\frac{\sigma_X^2}{n_X}+\frac{\sigma_Y^2}{n_Y}}} \underset{H_0}{\sim} \mathcal{N}(0,1)$
	large	unknown	$\dfrac{(\overline{X}-\overline{Y})-\delta}{\sqrt{\frac{s_X^2}{n_X}+\frac{s_Y^2}{n_Y}}} \underset{H_0}{\sim} \mathcal{N}(0,1)$
	small	unknown $\sigma_X = \sigma_Y$	$\dfrac{(\overline{X}-\overline{Y})-\delta}{s\sqrt{\frac{1}{n_X}+\frac{1}{n_Y}}} \underset{H_0}{\sim} t_{n_X+n_Y-2}$
Normal, paired $D_i = X_i - Y_i$	any $n_X = n_Y$	unknown	$\dfrac{\overline{D}-\delta}{\frac{s_D}{\sqrt{n}}} \underset{H_0}{\sim} t_{n-1}$

Région d'acceptation pour T normale : il faut vérifier à quelle zone correspond la valeur t.



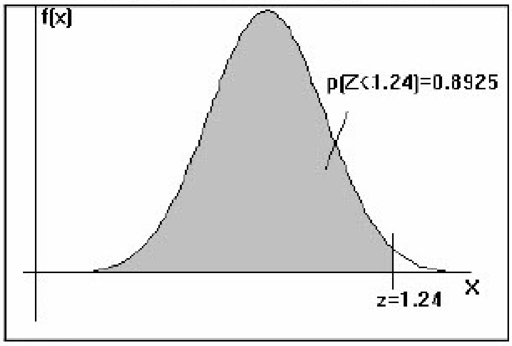
Statistique pour les tests.

	Parameter	Statistic	Null Hypothesis in Hypothesis Test	Standard Error	Test Statistic for Hypothesis Test	z or t ?	Degrees of Freedom	Conditions for Inference
Mean of a Population (known standard deviation)	μ	$\bar{x}$	$\mu = \mu_0$	$\dfrac{\sigma}{\sqrt{n}}$	$\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}$	z		σ known
Mean of a Population (unknown standard deviation)	μ	$\bar{x}$	$\mu = \mu_0$	$\dfrac{s}{\sqrt{n}}$	$\dfrac{\bar{x}-\mu}{s/\sqrt{n}}$	t	$n-1$	$n < 15$ + Normally dist. pop.: OK $n < 15$ + skewed dist./outliers.: not OK $n \geq 15$ + strong skew./ extreme outliers.: not OK $n \geq 40$: OK

	Parameter	Statistic	Null Hypothesis in Hypothesis Test	Standard Error for Confidence Interval	Standard Error for Hypothesis Test	Test Statistic for Hypothesis Test	z or t ?	Conditions for Inference
Proportion of a Population	p	$\hat{p}$	$p = p_0$	$\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$	$\sqrt{\dfrac{p_0(1-p_0)}{n}}$	$\dfrac{\hat{p}-p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$	z	CI: at least 15 successes and 15 failures HT: $np_0 \geq 10$, $n(1-p_0) \geq 10$
Proportion of Two Populations	$p_1 - p_2$	$\hat{p}_1 - \hat{p}_2$	$p_1 = p_2$	$\sqrt{\dfrac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \dfrac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$	$\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)}$	$\dfrac{\hat{p}_1-\hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)}}$	z	CI: at least 10 successes and 10 failures in each sample HT: at least 5 successes and 5 failures in each sample

TABLE DE LA LOI NORMALE CENTREE REDUITE

Lecture de la table: Pour z=1.24 (intersection de la ligne 1.2 et de la colonne 0.04), on a la proportion P(Z < 1,24) = 0.8925



$P(Z > 1{,}96) = 0{,}025$
$P(Z > 2{,}58) = 0{,}005$
$P(Z > 3{,}29) = 0{,}0005$

Rappels:

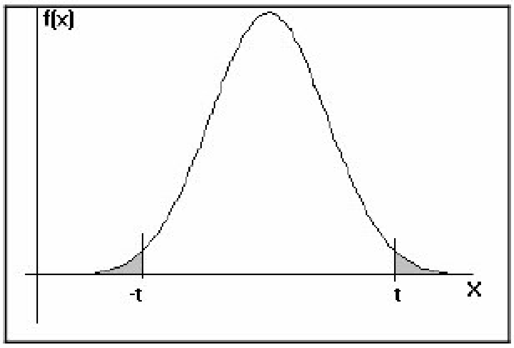
1/ $P(Z > z) = 1 - P(Z < z)$ et 2/ $P(Z < -z) = P(Z > z)$
Exemple: Sachant $P(Z < 1{,}24) = 0{,}8925$, on en déduit:
1/ $(P(Z > 1{,}24) = 1 - P(Z < 1{,}24) = 1 - 0{,}8925 = 0{,}1075$
2/ $P(Z < -1{,}24) = P(Z > 1{,}24) = 0{,}1075$

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,99865	0,99869	0,99874	0,99878	0,99882	0,99886	0,99889	0,99893	0,99896	0,99900
3,1	0,99903	0,99906	0,99910	0,99913	0,99916	0,99918	0,99921	0,99924	0,99926	0,99929
3,2	0,99931	0,99934	0,99936	0,99938	0,99940	0,99942	0,99944	0,99946	0,99948	0,99950
3,3	0,99952	0,99953	0,99955	0,99957	0,99958	0,99960	0,99961	0,99962	0,99964	0,99965
3,4	0,99966	0,99968	0,99969	0,99970	0,99971	0,99972	0,99973	0,99974	0,99975	0,99976
3,5	0,99977	0,99978	0,99978	0,99979	0,99980	0,99981	0,99981	0,99982	0,99983	0,99983
3,6	0,99984	0,99985	0,99985	0,99986	0,99986	0,99987	0,99987	0,99988	0,99988	0,99989
3,7	0,99989	0,99990	0,99990	0,99990	0,99991	0,99991	0,99992	0,99992	0,99992	0,99992
3,8	0,99993	0,99993	0,99993	0,99994	0,99994	0,99994	0,99994	0,99995	0,99995	0,99995
3,9	0,99995	0,99995	0,99996	0,99996	0,99996	0,99996	0,99996	0,99996	0,99997	0,99997
4,0	0,99997	0,99997	0,99997	0,99997	0,99997	0,99997	0,99998	0,99998	0,99998	0,99998

DISTRIBUTIONS DU t DE STUDENT

Table des valeurs critiques bilatérales usuelles

Pour une distribution de Student à ddl degrés de liberté et pour une proportion α (.05, .01 ou .001), la table indique t tel que $P(|T| > t) = \alpha$



Exemple: Pour ddl = 5, on a $P(|T| > 2.571) = .05$ (on note $t_{[5].05}$ cette valeur.).

α α/2 ddl	0,05 0,025	0,01 0,005	0,001 0,0005
1	12.706	63.657	636.619
2	4.303	9.925	31.599
3	3.182	5.841	12.924
4	2.776	4.604	8.610
5	2.571	4.032	6.869
6	2.447	3.707	5.959
7	2.365	3.499	5.408
8	2.306	3.355	5.041
9	2.262	3.250	4.781
10	2.228	3.169	4.587
11	2.201	3.106	4.437
12	2.179	3.055	4.318
13	2.160	3.012	4.221
14	2.145	2.977	4.140
15	2.131	2.947	4.073
16	2.120	2.921	4.015
17	2.110	2.898	3.965
18	2.101	2.878	3.922
19	2.093	2.861	3.883
20	2.086	2.845	3.850
21	2.080	2.831	3.819
22	2.074	2.819	3.792
23	2.069	2.807	3.768
24	2.064	2.797	3.745
25	2.060	2.787	3.725
26	2.056	2.779	3.707
27	2.052	2.771	3.690
28	2.048	2.763	3.674
29	2.045	2.756	3.659
30	2.042	2.750	3.646
40	2.021	2.704	3.551
60	2.000	2.660	3.460
120	1.980	2.617	3.373
30000	1.960	2.576	3.291

TABLE TEST Z

α	0,05	0,01	0,001
α/2	0,025	0,005	0,0005
Z	1,96	2,58	3,29

DISTRIBUTION DU KHI2

La table donne les valeurs critiques de χ^2 pour un nombre de degrés de liberté (ddl) et pour un seuil repère donnés (α).

Par exemple:

Pour ddl = 3 et α= 0,05 la table indique $\chi^2 = 7{,}81$

Ceci signifie que: $P(\chi^2_{[3]} > 7{,}81) = 0{,}05$

ddl \ α	0,05	0,01	0,001
1	3,84	6,63	10,83
2	5,99	9,21	13,82
3	7,81	11,34	16,27
4	9,49	13,28	18,47
5	11,07	15,09	20,52
6	12,59	16,81	22,46
7	14,07	18,48	24,32
8	15,51	20,09	26,12
9	16,92	21,67	27,88
10	18,31	23,21	29,59
11	19,68	24,72	31,26
12	21,03	26,22	32,91
13	22,36	27,69	34,53
14	23,68	29,14	36,12
15	25,00	30,58	37,70
16	26,30	32,00	39,25
17	27,59	33,41	40,79
18	28,87	34,81	42,31
19	30,14	36,19	43,82
20	31,41	37,57	45,31
21	32,67	38,93	46,80
22	33,92	40,29	48,27
23	35,17	41,64	49,73
24	36,42	42,98	51,18
25	37,65	44,31	52,62
26	38,89	45,64	54,05
27	40,11	46,96	55,48
28	41,34	48,28	56,89
29	42,56	49,59	58,30
30	43,77	50,89	59,70