université de BORDEAUX

PASS

TUTORAT SANTÉ BORDEAUX

Préparation aux examens Médicaux et Paramédicaux

Kinésithérapie
Ergothérapie
Psychomotricité
Podologie

Médecine Pharmacie Maïeutique Odontologie Filières Paramédicales

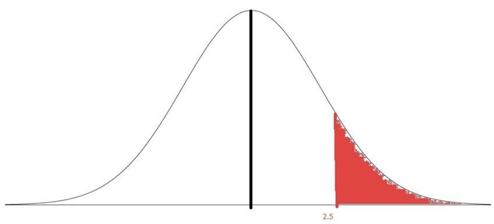
CORRECTION COLLE n°3 - UE5

12/10/2020 - Fait par la séance du jeudi

QCM 1 : BCE

On appliquera toujours la même méthode pour résoudre ce genre d'exercices. D'abord un changement de variable, puis un raisonnement classique sur la loi normale centrée réduite *(en s'aidant des schémas)*.

A. FAUX, On cherche P(X > 70).



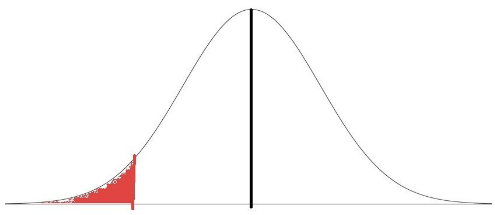
Changement de variable : $U_\alpha = \frac{X-\mu}{\sigma} = \frac{70-60}{4} = \frac{10}{4} =$ **2,5**.

On cherche dans la TER le α correspondant au U_α = 2,5.

On trouve un α = **0,01**.

A l'aide du graphique, on détermine la formule à utiliser : $P(X > 70) =$
$P(\frac{X-\mu}{\sigma} > \frac{70-60}{4} = 2,5) = \frac{\alpha}{2} = \frac{0,01}{2} =$ **0,005**.

B. VRAI, On cherche P(X < 58).



Changement de variable : $U_\alpha = \frac{X-\mu}{\sigma} = \frac{58-60}{4} = \frac{-2}{4} =$ **- 0,5**.

On cherche dans la TER le α correspondant au U_α = - 0,5.

On trouve α = **0,62**.

A l'aide du graphique, on détermine la formule à utiliser :
$P(U_\alpha < -0,5) = \frac{\alpha}{2} = \frac{0,62}{2} =$ **0,31**.

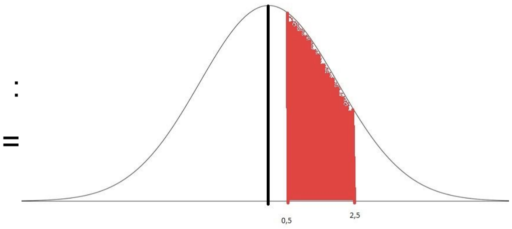
C. VRAI, On cherche P(62 <X < 70).

Changement de variable : $U_{\alpha_1} = \frac{X-\mu}{\sigma} = \frac{62-60}{4} = \frac{2}{4} =$ **0,5** ET $U_{\alpha_2} = \frac{X-\mu}{\sigma} = \frac{70-60}{4} = \frac{10}{4} =$ **2,5**.

On cherche dans la TER le α_1 correspondant au U_{α_1} = 0,5 ET le α_2 correspondant au U_{α_2} = 2,5.

On trouve α_1 = **0,62** et α_2 = **0,01**.

A l'aide du graphique, on détermine la formule à utiliser : $P(0,5 < U < 2,5) = \frac{\alpha_1}{2} - \frac{\alpha_2}{2} = \frac{0,62}{2} - \frac{0,01}{2} =$ 0,31 - 0,005 = **0,305**.



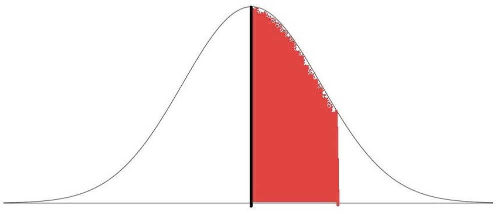
D. FAUX, On cherche P(60 < X < 64).

Changement de variable : $U_{\alpha_1} = \frac{X-\mu}{\sigma} = \frac{60-60}{4} =$ **0** ET $U_{\alpha_2} = \frac{X-\mu}{\sigma} = \frac{64-60}{4} = \frac{4}{4} =$ **1**.

On cherche dans la TER le α_2 correspondant au U_{α_2} = 1.

On trouve un α_2 = **0,32**.

A l'aide du graphique, on détermine la formule à utiliser : $P(0 < U_\alpha < 1) = \frac{1}{2} - \frac{\alpha_2}{2} = 0,5 - \frac{0,32}{2} =$ 0,5 - 0,16 = **0,34**.



E. VRAI, On cherche P(56 < X < 64).

Changement de variable : $U_{\alpha_1} = \frac{X-\mu}{\sigma} = \frac{56-60}{4} = \frac{-4}{4} =$ **-1** ET $U_{\alpha_2} = \frac{X-\mu}{\sigma} = \frac{64-60}{4} = \frac{4}{4} =$ **1**.

On cherche donc P(-1<U<1) ce qui correspond à P($\mu - \sigma$ < U < $\mu + \sigma$) = **68%**.

Rappel : ***À CONNAÎTRE PAR COEUR*** *(cf. diapo Probabilités p.64/65)*

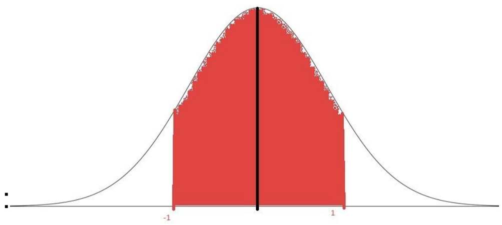
- **P($\mu - \sigma$ < U < $\mu + \sigma$) = 68%.**
- **P($\mu - 2\sigma$ < U < $\mu + 2\sigma$) = 95%.**

On peut quand même appliquer la méthode habituelle :

On cherche dans la TER le α correspondant au U_α = -1 et 1.

On trouve α = **0,32**.

A l'aide du graphique, on détermine la formule à utiliser : $P(-1 < U_\alpha < 1) = 1 - \alpha = 1 - 0,32 =$ 0,68 soit 68%.



QCM 2 : AC

A. VRAI, Ici, on a le paramètre (f) de l'échantillon connu et celui de la population (P) inconnu. On réalise bien un **intervalle de confiance** pour estimer la fréquence théorique P dans la population à partir des données de l'échantillon.

B. FAUX, Dans le cas de la réalisation d'un intervalle de confiance pour des variables qualitatives, les conditions sont à vérifier à **POSTERIORI**.

Remarque : On peut vérifier à priori que Nf et N(1-f) $\geq$ 5 car si Nf et N(1-f) < 5, NP et NQ seront inférieurs à 5 également et nous n'aurons pas perdu de temps à calculer un intervalle où les conditions ne sont pas respectées. La réciproque n'est pas vrai, donc si Nf et N(1-f) $\geq$ 5 à priori, il faudra tout de même vérifier que NP et NQ $\geq$ 5 à posteriori.

Ici, Nf = 400 $\times$ 0,15 = 60 et N(1-f) = 400 $\times$ 0,85 = 340. Nf et N(1-f) $\geq$ 5 donc nous allons calculer cet intervalle de confiance, mais il ne faudra pas oublier de vérifier à posteriori que NP et NQ $\geq$ 5.

P $\in [f \pm U\alpha\sqrt{\frac{f(1-f)}{N}}]$ ⇔ P $\in [0,15 \pm 2\sqrt{\frac{0,15(1-0,15)}{400}}]$ ⇔ P $\in [0,15 \pm 2\sqrt{\frac{0,15 \times 0,85}{400}}]$ ⇔ P $\in [0,15 \pm 2\sqrt{\frac{1275 \times 10^{-4}}{400}}]$ ⇔ P $\in [0,15 \pm 2 \times \frac{36 \times 10^{-2}}{20}]$ ⇔ P $\in [0,15 \pm \frac{36 \times 10^{-2}}{10}]$ ⇔ P $\in [0,15 \pm 36 \times 10^{-3}]$ ⇔ P $\in [0,15 \pm 0,036]$ ⇔ **P ∈ [0,114 ; 0,186]**.

On peut maintenant vérifier les conditions de validité : NP = $400 \times 0,114 = 45,6$ et NQ = $400 \times (1 - 0,186) = 325,6$. Les conditions sont bien respectées.

C. VRAI, *cf B.*

D. FAUX, Ici, Nf = 64 $\times$ 0,15 = 9,6 et N(1-f) = 64 $\times$ 0,85 = 54,4. Nf et N(1-f) $\geq$ 5 donc nous allons calculer cet intervalle de confiance, mais il ne faudra pas oublier de vérifier à posteriori que NP et NQ $\geq$ 5.

P $\in$ $[f \pm U\alpha\sqrt{\frac{f(1-f)}{N}}]$ $\Leftrightarrow$ P $\in$ $[0,15 \pm 2\sqrt{\frac{0,15(1-0,15)}{64}}]$ $\Leftrightarrow$ P $\in$ $[0,15 \pm 2\sqrt{\frac{0,15\times0,85}{64}}]$ $\Leftrightarrow$ P $\in$ $[0,15 \pm 2\sqrt{\frac{1275\times10^{-4}}{64}}]$ $\Leftrightarrow$ P $\in$ $[0,15 \pm 2 \times \frac{36\times10^{-2}}{8}]$ $\Leftrightarrow$ P $\in$ $[0,15 \pm \frac{2\times9\times4\times10^{-2}}{4\times2}]$ $\Leftrightarrow$ P $\in$ $[0,15 \pm 9 \times 10^{-2}]$ $\Leftrightarrow$ P $\in$ $[0,15 \pm 0,09]$ $\Leftrightarrow$ P $\in$ [0,06 ; 0,24].

On peut maintenant vérifier les conditions de validité : NP = $64 \times 0,06 = 3,84$ donc NP < 5. ***ATTENTION***, **LES CONDITIONS DE VALIDITÉ NE SONT PAS RESPECTÉES !** L'intervalle calculé est donc faux.

E. FAUX, Le temps de révision des étudiant(e)s n'a strictement rien à voir avec l'énoncé. De plus, l'item est doublement faux puisqu'il ne fait pas mention de risque dans sa formulation.

QCM 3 : ABD

A. VRAI.

B. VRAI, Ici, on effectue un test de comparaison de 2 moyennes observées dans le cas d'échantillons petits (car $N_2 < 30$) et indépendants. On va donc procéder au test de Student sachant que les conditions sont vérifiées :

- N1 et/ou N2 < 30
- échantillons indépendants
- les variances ne sont pas significativement différentes
- les variables suivent la loi normale

$DDL = N_1 + N_2 - 2 = 50 + 16 - 2 = \mathbf{64}$.

Remarque : La DDL est le dénominateur dans le calcul de la variance commune.

C. FAUX, $s^2 = \frac{s_1^2 \, x \, N_1 + s_2^2 \, x \, N_2}{N_1 + N_2 - 2} = \frac{8 \times 50 + 11 \times 16}{50 + 16 - 2} = \frac{400 + 176}{64} = \frac{576}{64} = \mathbf{9}$.

$$T = \frac{m_1 - m_2}{s\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}} = \frac{28 - 25}{\sqrt{9}\sqrt{\frac{1}{50} + \frac{1}{16}}} = \frac{3}{3\sqrt{\frac{4}{50}}} = \frac{1}{\frac{2}{7}} = \frac{7}{2} = \mathbf{3,5}.$$

D. VRAI, Dans notre table de Student, on ne trouve pas le T_α exact pour un DDL de 64 et un α de 5% mais on voit tout de même que, en se plaçant entre une DDL de 40 et 80, la RC se trouve entre]-∞ ; - 1,99] $\cup$ [+ 1,99 ; +∞[et]-∞ ; - 2,021] $\cup$ [+ 2,021 ; +∞[.

Dans ces 2 cas, $3,5 \in RC$, donc on **rejette H0 au risque α de première espèce de se tromper**.

Remarque : Quand vous voyez que pour un DDL demandé vous ne trouvez pas la valeur exacte du T_α, il ne faut pas paniquer et vous bloquer. On ne va pas vous piéger là-dessus donc selon que vous prendrez la borne inférieure ou supérieure, ça ne changera pas la conclusion.

E. FAUX, On rejette H0 au risque α donc il y a une différence significative entre les moyennes étudiées de ces deux échantillons, ce qui signifie qu'ils ne sont pas issus de la même population d'étudiants.

QCM 4 : C

A. FAUX, C'est **H0**. **H1** correspond à **l'hypothèse inverse/alternative** de H0 et serait "Il y a une différence significative" ou "L'échantillon n'est pas représentatif de la population".

B. FAUX, On cherche à comparer une fréquence théorique et une fréquence observée : on va utiliser le test basé sur la loi normale.

Les conditions NP et NQ $\geq$ 5 sont bien respectées (NP = 144 x 0,2 = 28,8 et NQ = 144 x 0,8 = 115,2).

Cependant, la formule à appliquer est U = $\dfrac{f-P}{\sqrt{\frac{P(1-P)}{N}}}$.

C. VRAI, U = $\dfrac{0{,}26-0{,}2}{\sqrt{\frac{0{,}2\times0{,}8}{144}}} = \dfrac{0{,}06}{\sqrt{\frac{0{,}16}{144}}} = \dfrac{0{,}06}{\frac{\sqrt{16\times10^{-2}}}{\sqrt{144}}} = \dfrac{0{,}06}{\frac{4\times10^{-1}}{12}} = \dfrac{0{,}06\times12}{4\times10^{-1}} = \dfrac{0{,}06\times4\times3}{4\times10^{-1}}$

$= 0,6 \times 3 =$ **1,8**.

D. FAUX, Pour α = 1%, on trouve dans la TER que U_α = 2,576 : la région critique est]-∞ ; - 2,576] ∪ [+ 2,576 ; +∞[donc U ∉ RC, on accepte l'hypothèse nulle H0 donc il n'y a pas de différence significative entre l'échantillon et la population au **risque β** (**et non ALPHA**, sorry).

E. FAUX, Il n'y a pas de différence significative entre l'échantillon et la population donc **l'échantillon est représentatif de la population au risque β**.

QCM 5 : DE

A. FAUX, Le **coefficient de corrélation** est la mesure de l'association entre deux variables **QUANTITATIVES**. Elle permet de préciser la **force** (ou l'intensité) **de la liaison**, s'il y en a une.

B. FAUX, La **régression** est l'étude de l'association entre deux variables **QUANTITATIVES**. Elle permet de **prédire l'évolution d'une variable en fonction de celle à laquelle elle est liée**.

C. FAUX, L'information donnée est **QUALITATIVE**. Par exemple : si les points sont alignés (forme allongée du graphique), une liaison entre les deux variables est **possible**.

D. VRAI, Ne pas oublier que le **coefficient de corrélation (r)** donne **2 informations** : la **force** et le **sens de l'association linéaire** entre les deux variables étudiées.

E. VRAI, En effet, le **coefficient de corrélation r** est **lié** à la **pente a de la droite de régression** par la relation : $r = a\frac{s_X}{s_Y}$.

Par exemple : S'il n'y a pas de corrélation entre les deux variables, c'est-à-dire que $r = 0$, alors la pente $a = 0$ → La droite de régression $y = ax + b$ devient $y = b$ (droite HORIZONTALE).

↳ **r et a sont donc bien proportionnels.**

QCM 6 : BD

A. FAUX, $r = \frac{Sxy}{Sx \times Sy}$ = $\frac{40}{20\times5}$ = $\frac{2}{5}$ = **0,4**.

B. VRAI, *cf A.*

C. FAUX, H0 : "Il y a indépendance linéaire" ou ""Il n'y a **pas** de corrélation entre les deux variables" ou "p=0".

D. VRAI, Tobs = $\frac{r}{\sqrt{1-r^2}}\sqrt{N-2} = \frac{0{,}4}{\sqrt{1-0{,}4^2}}\sqrt{38-2} = \frac{0{,}4}{\sqrt{1-0{,}16}}\sqrt{36} = \frac{0{,}4}{\sqrt{0{,}84}} \times 6 = \frac{0{,}4}{1} \times 6 =$ **2,4**.

Comme **N>30**, **T suit une loi normale** centrée réduite et donc T α est lu dans la **TER**.

Pour $\alpha = 5\%$, $T\alpha = 1,96$ donc **Tobs>T α** alors Tobs appartient à la région critique : on **rejette H0 avec un risque $\leq \alpha$**. On peut donc conclure qu'il y a **corrélation linéaire** entre les deux variables.

E. FAUX, ***ATTENTION*** : Ce test d'indépendance linéaire met en évidence une relation statistique mais **PAS un lien de causalité**.

QCM 7 : C

A. FAUX, La formule pour trouver la pente est : $a = \frac{Sxy}{Sx^2}$. ***ATTENTION*** à ne pas oublier le **carré** !!!

B. FAUX, Pour calculer l'équation de la droite de régression, on calcule en premier lieu la pente a : $a = \frac{Sxy}{Sx^2}$ = $\frac{40}{20^2}$ = $\frac{2\times20}{20\times20} = \frac{2}{20} = \frac{1}{10} =$ **0,1**.

Ensuite on calcule b : $b = \overline{y} - a \times \overline{x} = 15 - 0,1 \times 225 = 15 - 22,5 =$ **- 7,5**.

On complète la formule $y = ax + b$ et on trouve $y = 0,1x - 7,5$.

C. VRAI, *cf B.*

D.FAUX, On reprend l'équation de la droite de régression que l'on complète : $y = 0,1x - 7,5 = 0,1 \times 210 - 7,5 = 21 - 7,5 = 13,5$ **.**

ATTENTION AUX UNITÉS !! Le résultat est 13,5 **mm** !!

E. FAUX, L'étude de la corrélation est réalisée pour des quantités de jus d'asperge comprises entre **200 et 250 mL**. On ne peut donc pas utiliser l'équation de la droite trouvée pour une quantité de 150mL, car cette droite de régression n'est valable qu'entre 200 et 250mL.

QCM 8 : B

A. FAUX, Ici, si les conditions sont respectées, on réalise un **Chi 2 d'indépendance** et non pas d'ajustement. En effet, dans le cas d'un Chi 2 d'ajustement, on compare une fréquence observée avec une fréquence théorique. Or, dans notre étude on réalise une **comparaison entre deux fréquences observées** (fréquence d'énigmite chez les tuteurs d'UE Noire et fréquence d'énigmite chez les autres tuteurs).

On peut dire aussi qu'on teste le lien (indépendance) entre les 2 variables énigmite et UE.

Vérifions les conditions :

On construit le tableau des **effectifs observés.**

TCO	Énigmite	~~Énigmite~~	TOTAL
UE Noire	26	4	**30**
~~UE Noire~~	**14** *(= 40 × 0,35)*	156	**170**
TOTAL	**40**	160	200

Sous H0, on en déduit le tableau des **effectifs théoriques.**

TCT	Énigmite	~~Énigmite~~	TOTAL
UE Noire	$\frac{40 \times 30}{200}$ = 6	24	30
~~UE Noire~~	34	136	170
TOTAL	40	160	200

Tous les effectifs théoriques étant ≥ 5, les conditions sont respectées : on peut réaliser un Khi 2 d'indépendance.

B. VRAI.

C. FAUX, ***ATTENTION*** : ce sont les **effectifs THÉORIQUES qui doivent être ≥ 5** et non pas les effectifs observés. Ainsi, si un effectif observé est < 5 (comme dans notre étude où un des effectifs observés est de 4), on peut quand même réaliser le test si tous les Et $\geq$ 5.

D. FAUX, ***ATTENTION*** **:** le but d'un Chi 2 d'indépendance est de tester l'indépendance entre **2 VARIABLES**, et non pas echantillons.

E. FAUX, Le test **ANOVA** est utilisé dans le cadre de **variables quantitatives** et non pas qualitatives comme le test du Khi2 : il n'est donc pas réalisable ici. ***ATTENTION*** aux items comme celui-ci qui peuvent perturber de par leur incohérence.

QCM 9 : BE

A. FAUX, **DDL = (ligne - 1) x (colonne - 1)** = (2 - 1) x (2-1) = **1**.

B. VRAI, Avec un DDL de 1 et α = 5%, on lit dans la table du χ^2 un χ^2_α = **3,84**. Sachant que la RC = [χ^2_α; + ∞[, on en déduit que la RC correspond bien à **[3,84 ; + ∞[**.

C. FAUX, $\chi^2 = \Sigma\frac{(Eo-Et)^2}{Et} = \frac{(26-6)^2}{6} + \frac{(4-24)^2}{24} + \frac{(14-34)^2}{34} + \frac{(156-136)^2}{136} = \frac{20^2}{6} + \frac{20^2}{24} + \frac{20^2}{34} + \frac{20^2}{136} \simeq$

66 + 16 + 11 + 2 = **95**. Donc X² ∈ RC.

Astuce : Pour gagner du temps, on remarque que $\frac{20^2}{6}$ est bien supérieur à 3,84 donc inutile de continuer le calcul.

Remarque : Lorsque **X² ∈ RC**, on **rejette H0** au **risque α de première espèce ≤ 5%**. Les items C et D étaient donc faux et on pouvait les éliminer sans calcul.

D. FAUX, *cf C.*

E. VRAI, Ici, X² ∈ RC donc on rejette H0 : il y a bien un **lien** entre être un tuteur d'UE noire et être atteint d'énigmite, et la fréquence est supérieure dans le groupe UE Noire. On conclut au risque α de première espèce ≤ 5%.

QCM 10 : BCE

A. FAUX, Ici, on compare une **fréquence observée à une fréquence théorique** donc il ne peut pas s'agir d'un Khi² d'indépendance mais d'un **Khi² d'ajustement**.

B. VRAI, *cf. A.*

C. VRAI, En appliquant la fréquence de la population à l'échantillon on obtiendra les effectifs théoriques.

D. FAUX, On construit le Tableau Observé à partir des données de l'énoncé :

TO	Boussole	Carte routière	GPS	Effectif total
Eo	60	90	50	200

Sous H0, on réalise le Tableau Théorique et on applique à notre échantillon de 200 les fréquences théoriques :

TT	Boussole	Carte routière	GPS	Effectif total
Et	50	40	110	200

On commence par vérifier que tous les **effectifs théoriques soient** ≥ 5, ce qui est le cas ici.

De plus, on a 3 colonnes donc **2 DDL** car DDL = (nombre de colonnes - 1) avec un risque α précisé dans l'énoncé de **1%**.

En lisant la table du Khi², on trouve que la région critique se situe dans l'intervalle **[9,21 ; + ∞ [**.

On obtient donc :

$\chi^2 = \Sigma\frac{(Eo-Et)^2}{Et} = \frac{(60-50)^2}{50} + \frac{(90-40)^2}{40} + \frac{(50-110)^2}{110} = \frac{100}{50} + \frac{2500}{40} + \frac{3600}{110} \simeq 2 + 62,5 + 32,7 =$ **97,2**.

Le **paramètre appartient donc à la région critique**.

Astuce :

- Pour trouver 32,7, au concours/colle vous arrondissez ++.
- On remarque qu'on peut s'arrêter au deuxième calcul car on obtient une valeur (62,5) qui est largement supérieure à 9,21, donc le paramètre sera forcément dans la région critique.

E. VRAI.