

Correction du CC 2025/2026
Proba - VA - Intervalles de confiance

Septembre 2026

QCM 12 - Énoncé

Parmi les propositions suivantes, indiquez la ou les réponse(s) juste(s) :

- A. on calcule un intervalle de confiance de l'estimation de la moyenne
- B. un intervalle de fluctuation est centré sur m , l'estimation de la moyenne
- C. soient A et B 2 évènements non vides, $\overline{A \cap B} = \overline{A} \cap \overline{B}$
- D. une variable aléatoire discrète prend un nombre fini ou infini dénombrable de valeurs possibles
- E. soit X une variable aléatoire continue, $P(a \leq X \leq b)$ correspond à l'aire sous la courbe représentant la fonction de répartition de X entre les bornes a et b

QCM 12 - Réponses

Item A : on calcule un intervalle de confiance de l'estimation de la moyenne

On calcule un intervalle de confiance d'un paramètre **théorique** :

$$ic_{1-\alpha}(\mu)$$

Item A : Faux

Item B : un intervalle de fluctuation est centré sur m , l'estimation de la moyenne

Un intervalle de fluctuation est centré sur un paramètre théorique (μ ou π , cf diapo 52)

Item B : Faux

Item C : soient A et B 2 évènements non vides, $\overline{A \cap B} = \overline{A} \cap \overline{B}$
 $\overline{A \cap B} = \overline{A} \cup \overline{B}$ (cf diapo 7)

Item C : Faux

QCM 12 - Réponses

Item D : une variable aléatoire discrète prend un nombre fini ou infini dénombrable de valeurs possibles

cf diapo 34

Item D : Vrai

Item E : soit X une variable aléatoire continue, $P(a \leq X \leq b)$ correspond à l'aire sous la courbe représentant la fonction de répartition de X entre les bornes a et b

$P(a \leq X \leq b)$ correspond à l'aire sous la courbe représentant la fonction densité de probabilité de X (cf diapo 40)

Item E : Faux

QCM 13 - Énoncé

L'IMC correspond à l'indice de masse corporelle défini par le rapport entre le poids d'un individu (en kg) et le carré de sa taille (en m). Lorsque l'IMC d'un individu est supérieur à 30, cet individu est considéré comme obèse.

On considère les 2 variables aléatoires suivante :

- ▶ X_F , qui modélise l'IMC des femmes françaises en 2009. X_F suit une loi normale de paramètres $\mu_F = 23,2$ kg/m² et $\sigma_F = 5$ kg/m²
- ▶ X_H , qui modélise l'IMC des hommes français en 2009. X_H suit une loi normale de paramètres $\mu_H = 24,6$ kg/m² et $\sigma_H = 7$ kg/m²

On définit alors :

- ▶ la variable $X_T = \frac{X_F + X_H}{2}$.
- ▶ la variable Y_{HO} qui modélise le fait qu'un homme soit obèse ou non
- ▶ la variable W_{HO} qui modélise le nombre d'hommes obèses dans un échantillon de 20 hommes

Pour les calculs, conserver autant de chiffres significatifs que possible.
Parmi les propositions suivantes, indiquez la ou les réponse(s) juste(s)

QCM 13 - Réponse

Informations de l'énoncé

- ▶ $X_F \rightarrow \mathcal{N}(23,2; 5)$
- ▶ $X_H \rightarrow \mathcal{N}(24,6; 7)$
- ▶ $X_T = \frac{X_F + X_H}{2}$

Item A : X_T suit une loi normale de paramètres : $\mu_T = 23,9 \text{ kg/m}^2$
et $\sigma_T = \sqrt{37}$

X_T est une combinaison linéaire de 2 variables qui suivent des lois normales, donc X_T suit une loi normale

$$E(X_T) = \frac{1}{2}(E(X_F) + E(X_H)) = \frac{1}{2}(23,2 + 24,6) = 23,9$$

$$\text{var}(X_T) = \left(\frac{1}{2}\right)^2 \times (\text{var}(X_F) + \text{var}(X_H)) = \frac{1}{4}(5^2 + 7^2) = 18,5$$

$$\sigma_{X_T} = \sqrt{\text{var}(X_T)} = \sqrt{18,5}$$

Item A : Faux

QCM 13 - Réponse

Informations de l'énoncé

- ▶ $X_H \rightarrow \mathcal{N}(24,6; 7)$
- ▶ Y_{HO} modélise le fait qu'un homme soit obèse ($\text{IMC} \geq 30$) ou non

Item B : $P(X_H \geq 30) \simeq 0,22$

$$P(X_H \geq 30) = P\left(Z \geq \frac{30-24,6}{\sqrt{7}}\right) = P(Z \geq 0,7714) \text{ (centrer et réduire)}$$

$$P(X_H \geq 30) = 1 - P(Z \leq 0,77)$$

On lit $P(Z \leq 0,77)$ dans la table de la fdr de la loi $\mathcal{N}(0,1)$

$$P(X_H \geq 30) = 1 - 0,7794 \simeq 0,22$$

Item B : Vrai

Item C : la variable aléatoire Y_{HO} suit une loi $\mathcal{B}(n = 20; p = 0,22)$

Y_{HO} est une variable qui ne prend que 2 valeurs possibles : 0 (homme non obèse) ou 1 (homme obèse)

Y_{HO} suit une loi de Bernoulli de paramètre

$$p = P(Y_{HO} = 1) = P(X_H \geq 30) \simeq 0,22$$

Item C : Faux

QCM 13 - Réponse

Informations de l'énoncé

- ▶ $X_H \rightarrow \mathcal{N}(24,6; 7)$
- ▶ W_{HO} modélise le nombre d'hommes obèses dans un échantillon de 20 hommes

Item D : La variable aléatoire W_{HO} suit une loi $Bern(p = 0,22)$

$W_{HO} = \sum_{i=1}^{20} Y_{HOi}$ où les variables Y_{HOi} sont toutes indépendantes et de même loi de Bernoulli.

$W_{HO} \rightarrow \mathcal{B}(n = 20; p = 0,22)$

Item D : Faux

Item E : La probabilité d'avoir au moins un individu obèse dans l'échantillon est supérieure à 99%

On cherche $P(W_{HO} \geq 1)$

$$P(W_{HO} \geq 1) = 1 - P(W_{HO} = 0) = 1 - C_{20}^0 (0,22)^0 (1 - 0,22)^{20}$$

$$P(W_{HO} \geq 1) = 1 - 1 \times 1 \times 0,78^{20} \geq 0,99$$

Item E : Vrai

QCM 14 - Énoncé

Une enquête a été réalisée sur les habitudes en matière de tri et de consommation des ménages. Cette enquête a montré que 60% des ménages indiquaient pratiquer régulièrement le tri sélectif. L'enquête a aussi montré que la probabilité de consommer des aliments issus de l'agriculture biologique était égale à 0,4 chez les ménages pratiquant régulièrement le tri sélectif, à 0,3 chez les ménages pratiquant occasionnellement le tri sélectif et à 0,2 chez les ménages ne pratiquant pas du tout le tri sélectif. Enfin, 21% des ménages pratiquent le tri sélectif occasionnellement mais ne consomment pas d'aliments issus de l'agriculture biologique.

Notation

On note

- ▶ B la variable modélisant le fait de consommer des aliments issus de l'agriculture biologique
- ▶ T_j , T_o et T_r le fait de pratiquer le tri sélectif respectivement jamais, occasionnellement et régulièrement. T_j , T_o et T_r forment un système complet d'évènement

QCM 14 - Réponse

Informations de l'énoncé

$$P(T_r) = 0,6, P(B|T_r) = 0,4, P(B|T_o) = 0,3, P(B|T_j) = 0,2, \\ P(T_o \cap \bar{B}) = 0,21$$

Item A : 30% des ménages pratiquent occasionnellement le tri sélectif

$$P(T_o) = P(T_o \cap B) + P(T_o \cap \bar{B}) \text{ (début formule des probas totales)}$$

$$P(T_o) = P(B|T_o) \times P(T_o) + 0,21 = 0,3 \times P(T_o) + 0,21$$

$$\Leftrightarrow P(T_o)(1 - 0,3) = 0,21$$

$$\Leftrightarrow P(T_o) = \frac{0,21}{0,7} = 0,3$$

Item A : Vrai

$$\text{On en déduit } P(T_j) = 1 - P(T_r) - P(T_o) = 0,1$$

QCM 14 - Réponse

Informations de l'énoncé

$$P(T_r) = 0,6, P(B|T_r) = 0,4, P(B|T_o) = 0,3, P(B|T_j) = 0,2, \\ P(T_o \cap \bar{B}) = 0,21$$

Item B : 35% des ménages consomment des aliments issus de l'agriculture biologique

$$P(B) = P(B|T_r)P(T_r) + P(B|T_o)P(T_o) + P(B|T_j)P(T_j) \text{ (formule des probas totales)}$$

$$P(B) = 0,4 \times 0,6 + 0,3 \times 0,3 + 0,2 \times 0,1 = 0,35$$

Item B : Vrai

Item C : 70% des ménages ne consommant pas d'aliments biologiques pratiquent régulièrement le tri sélectif

$$P(T_r|\bar{B}) = \frac{P(T_r \cap \bar{B})}{P(\bar{B})} = \frac{P(\bar{B}|T_r) \times P(T_r)}{P(\bar{B})} = \frac{(1-0,4) \times 0,6}{0,65} \simeq 0,55$$

Item C : Faux

QCM 14 - Réponse

Informations de l'énoncé

$$P(T_r) = 0,6, P(B|T_r) = 0,4, P(B|T_o) = 0,3, P(B|T_j) = 0,2, \\ P(T_o \cap \bar{B}) = 0,21$$

Item D : les évènements « pratiquer le tri sélectif » et « consommer des aliments issus de l'agriculture biologique » sont indépendants

$$P(T_r|\bar{B}) \simeq 0,55 \text{ et } P(T_r) = 0,6$$

Les 2 évènements ne sont pas indépendants

Item D : Faux

Item E : 95% des ménages pratiquent le tri sélectif de façon régulière ou consomment des aliments issus de l'agriculture biologique

$$P(T_r \cup B) = P(T_r) + P(B) - P(T_r \cap B)$$

$$P(T_r \cup B) = 0,6 + 0,35 - P(B|T_r) \times P(T_r) = 0,6 + 0,35 - 0,4 \times 0,6 = 0,71$$

Item E : Faux

QCM 15 - Énoncé

Afin d'évaluer la prise de poids moyenne lors de la prise d'un médicament M, un échantillon de 100 patients prenant le médicament M a été constitué. Les patients ont été suivis pendant 12 mois. A l'issue des 12 mois, la moyenne et l'écart-type de la prise de poids ont été estimés respectivement à 1,7 kg et 1,2 kg.

Les résultats sont donnés avec 1 chiffre après la virgule.

Parmi les propositions suivantes, indiquez la ou les réponse(s) juste(s).

- A. la borne supérieure de l'intervalle de confiance, à la confiance 0,95, de la prise de poids moyenne dans la population de patients prenant le médicament M vaut 1,9 kg
- B. la borne inférieure de l'intervalle de confiance, à la confiance 0,962, de la prise de poids moyenne dans la population de patients prenant le médicament M vaut 1,4 kg
- C. l'inclusion de 554 patients dans l'étude permet d'obtenir un intervalle de confiance à la confiance 0,95 ayant une précision égale à 100 g
- D. un ic contient toujours la valeur du paramètre à estimer
- E. la largeur de l'ic diminue quand la valeur du risque α augmente

QCM 15 - Réponse

Informations de l'énoncé

$$n = 100, m = 1,7 \text{ kg}, s = 1,2 \text{ kg}$$

Intervalle de confiance de la prise de poids **moyenne**

Item A : borne sup de l'ic à 95% de $\mu = 1,9 \text{ kg}$

Intervalle de confiance d'une moyenne (théorique), cas grand échantillon ($n=100$)

$$ic_{0,95}(\mu) = m \pm z_{1-\frac{\alpha}{2}} \times \frac{s}{\sqrt{n}}$$

$$m = 1,7 \quad s = 1,2 \quad z_{1-\frac{\alpha}{2}} = 1,96$$

Application numérique : $bs = 1,935$

Résultat avec un chiffre après la virgule (majoration) : $bs=2,0 \text{ kg}$

Item A : Faux

Item B : borne inf de l'ic à la confiance 0,962 de $\mu = 1,4 \text{ kg}$

$$\alpha = 1 - 0,962 = 0,038. \text{ On en déduit } 1 - \frac{\alpha}{2} = 0,981$$

On lit $z_{0,981}$ dans la table 2 et on obtient 2,0749

Application numérique : $binf = 1,451012$ On minore $binf$: $binf = 1,4$

Item B : Vrai

QCM 15 - Réponse

Informations de l'énoncé

$$n = 100, m = 1,7 \text{ kg}, s = 1,2 \text{ kg}$$

Intervalle de confiance de la prise de poids **moyenne**

Item C : l'inclusion de 554 patients dans l'étude permet d'obtenir un intervalle de confiance à la confiance 0,95 ayant une précision égale à 100 g

$$\text{Précision} = 1/2 \text{ largeur de l'ic} = z_{1-\frac{\alpha}{2}} \times \frac{s}{\sqrt{n}}$$

$$n = \frac{z_{1-\frac{\alpha}{2}}^2 \times s^2}{\text{precision}^2}$$

Attention à bien mettre tout dans la même unité (ici en kg)!

$$n = \frac{1,96^2 \times 1,2^2}{0,1^2} = 553,2$$

On devra donc bien inclure au moins 554 patients dans l'étude

Item C : Vrai

QCM 15 - Réponse

Item D : un ic contient toujours la valeur du paramètre à estimer

Parmi l'infinité d'ic possible, il y en a un pourcentage égal à α qui ne contient pas la vraie valeur du paramètre à estimer (cf diapo 37 du cours)

Item D : Faux

Item E : la largeur de l'ic diminue quand la valeur du risque α augmente

Diapo 43 du cours : la largeur de l'ic augmente avec le niveau de confiance.

La largeur de l'ic diminue quand le niveau de confiance diminue (ce qui correspond à une augmentation du risque α)

Item E : Vrai

Correction du CT 2025/2026
Proba - VA - Intervalles de confiance

Septembre 2026

QI 1 - Énoncé

Un biologiste s'intéresse à un nouveau score de prédiction de la pathogénicité des variants génétiques identifiés chez les patients. On note X la variable aléatoire qui modélise ce score et on suppose que X suit une loi normale d'espérance égale à 10 et d'écart-type égal à 2. Dans la publication décrivant ce nouveau score, il est indiqué que si le score est inférieur à 7, le variant est prédit bénin, si le score est supérieur à 13, le variant est prédit pathogène et si le score est compris entre 7 et 13, on ne peut pas conclure sur la base de ce score.

Vous ne ferez pas d'arrondi dans les calculs. Les résultats finaux seront arrondis avec 2 chiffres après la virgule.

- A. la valeur v du score telle que 75% des variants ont une valeur inférieure à v est environ égale à 11,35
- B. il y a entre 6% et 7% des variants qui sont prédits pathogènes avec ce score
- C. si un biologiste analyse 100 variants indépendants, la va modélisant le nb de variants prédits pathogènes avec ce score suit une loi de Bernoulli
- D. si un biologiste analyse 100 variants indépendants, la probabilité qu'au moins 5 d'entre eux soient prédits pathogènes vaut environ 0,75
- E. si un biologiste analyse 20 variants indépendants, la variable aléatoire modélisant la proportion de variants prédits pathogènes suit approximativement une loi normale d'espérance égale à 1,34 et d'écart-type égal à 1,12

Q1 - Correction

Informations de l'énoncé

- ▶ $X \rightarrow \mathcal{N}(10; 2)$
- ▶ Si score > 13 : variant prédit pathogène

Item A : valeur v telle que 75% des variants ont une valeur de score inférieure à v est environ égale à 11,35

On cherche v tel que $P(X \leq v) = 0,75$.

$$P(X \leq v) = P\left(\frac{X-10}{2} \leq \frac{v-10}{2}\right) = 0,75$$

$$P(X \leq v) = P(Z \leq z) = 0,75 \text{ avec } z = \frac{v-10}{2}$$

Dans la table 2, on lit la valeur z telle que $P(Z \leq z) = p$

Pour $p=0,75$, on lit dans la table $z = 0,6745$

Or, $v = 2 \times z + 10$, donc $v = 2 \times 0,6745 + 10 = 11,3490 \simeq 11,35$

Item A : Vrai

QI 1 - Correction

Informations de l'énoncé

- ▶ $X \rightarrow \mathcal{N}(10; 2)$ (X : score)
- ▶ Si score > 13 : variant prédit pathogène

Item B : il y a entre 6% et 7% des variants qui sont prédits pathogènes avec ce score

On cherche la proba qu'un variant soit pathogène : $P(X \geq 13)$

$$P(X \geq 13) = P\left(\frac{X-10}{2} \geq \frac{13-10}{2}\right) = P(Z \geq 1,5)$$

$$P(X \geq 13) = 1 - P(Z \leq 1,5) = 1 - \Phi(1,5) = 1 - 0,9332 = 0,0668$$

Item B : Vrai

QI 1 - Correction

Information des questions précédentes

Probabilité qu'un variant soit prédit pathogène = 0,067

Item C : si un biologiste analyse 100 variants indépendants, la va modélisant le nb de variants prédits pathogènes avec ce score suit une loi de Bernoulli

Soit Y la va modélisant le fait qu'un variant soit prédit pathogène

Y a 2 valeurs possibles :

- ▶ 0 si le variant n'est pas prédit pathogène (proba = $1-0,067$)
- ▶ 1 si le variant est prédit pathogène (proba = 0,067)

Y suit une loi $\text{Bern}(p=0,067)$

Soit Z la va modélisant le nb variants prédits pathogènes

$Z = \sum_{i=1}^{100} Y_i$, où les Y_i sont toutes indépendantes et de même loi de Bernoulli, $Z \rightarrow \mathcal{B}(n = 100; p = 0,067)$

Item C : Faux

QI 1 - Correction

Information des questions précédentes

- ▶ Z la va modélisant le nb variants prédits pathogènes
- ▶ $Z \rightarrow \mathcal{B}(n = 100; p = 0,067)$

Item D : si un biologiste analyse 100 variants indépendants, la proba qu'au moins 5 d'entre eux soient prédits pathogènes $\simeq 0,75$

On cherche $P(Z \geq 5)$. Afin de simplifier les calculs, on va vérifier si on peut approximer la loi binomiale suivie par Z par une loi normale

- ▶ $n = 100$ ($n \geq 30$)
- ▶ $n \times p = 100 \times 0,067 = 6,7$ ($np \geq 5$)
- ▶ $n \times (1 - p) = 100 \times 0,933 = 93,3$ ($n(1 - p) \geq 5$)

Les conditions d'approximation sont vérifiées, $Z \rightsquigarrow \mathcal{N}(6,7; 2,5)$

$P(Z \geq 5) = P\left(\frac{Z-6,7}{2,5} \geq \frac{5-6,7}{2,5}\right) = P(W \geq -0,68) = P(W \leq 0,68) \simeq 0,75$
avec $W \rightarrow \mathcal{N}(0; 1)$

Item D : Vrai

Q1 1 - Correction

Item E : si un biologiste analyse 20 variants indépendants, la variable aléatoire modélisant la proportion de variants prédits pathogènes suit approximativement une loi normale d'espérance égale à 1,34 et d'écart-type égal à 1,12

De la même façon qu'on avait défini Z , on définit R la variable aléatoire qui modélise le nombre de variants prédits pathogènes dans un ensemble de 20 variants.

$$R \rightarrow \mathcal{B}(n = 20, p = 0,067)$$

On vérifie si les conditions d'approximation de la loi binomiale suivie par R par une loi normale sont vérifiées :

- ▶ $n = 20 \quad (n \leq 30)$
- ▶ $n \times p = 20 \times 0,067 = 1,34 \quad (np \leq 5)$
- ▶ $n \times (1 - p) = 20 \times 0,933 = 18,66 \quad (n(1 - p) \geq 5)$

Les conditions ne sont pas vérifiées, on ne peut donc pas faire l'approximation

Item E : Faux