

Université Claude Bernard Lyon 1



Tutorat Lyon Est

Année Universitaire 2018 - 2019

Unité d'Enseignement 4

Concours de l'année 2017-2018

Correction détaillée

**Mathilde ABAD
Nona MANARYAN
Camille MARTIN**

Correction rapide

Questions	Item(s) juste(s)
1	BDE
2	BCDE
3	ABCE
4	AC
5	AE
6	BDE
7	AD
8	BCD
9	ADE
10	AC
11	ACD
12	ACD
13	AD
14	ANNULE

Correction détaillée

Question 1 : BDE

A FAUX Cette équation différentielle est bien du 2^{ème} ordre mais non linéaire, à coefficients constants et sans second membre.

Non linéaire par le terme « $4x \frac{dx}{dt}$ », à coefficients constants car aucun terme ne dépend de t et sans second membre car tous les termes de l'équation contiennent x ou sa dérivée.

B VRAI Cette équation ne contient pas de terme non linéaire, sa plus haute dérivée « $y^{(2)}$ » est d'ordre 2, elle possède 2 coefficients non constants qui sont $2x$ et $2x^2$, elle ne contient pas de second membre car tous les termes de l'équation contiennent y ou sa dérivée.

C FAUX L'équation de la dérivée peut s'écrire sous la forme : $B' = -2B$.

Ainsi en prenant en compte la condition initiale $B(0) = -3$ et la formule : $y' = a x y \Rightarrow y(t) = \lambda \cdot e^{at}$, on trouve : $B(t) = -3e^{-2t}$.

La croissance de notre équation dépend du signe de a et λ .

- a et λ de même signe : $a\lambda > 0$ donc fonction croissante.

- a et λ de signe opposé : $a\lambda < 0$ donc fonction décroissante.

Nous pouvons donc dire que notre fonction est croissante car a et λ sont de même signe, tous deux négatifs.

D VRAI De même que pour l'item C, on compare les signes de a et λ pour connaître la croissance de la fonction. Elle est donc décroissante car on a deux valeurs de signes opposés.

Elle est positive car une exponentielle est toujours positive et elle est multipliée par un nombre positif.

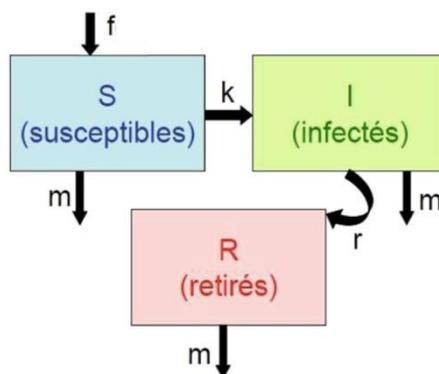
E VRAI L'équation : $\frac{dy}{dt} = \frac{y}{t}$ peut aussi s'écrire : $y' = \frac{1}{t} x y$

En prenant en compte la solution générale $y' = g(t) x y \Rightarrow y(t) = \lambda \cdot e^{G(t)}$. Or la primitive de $\frac{1}{t}$ est $\ln(t)$. Donc on se retrouve avec $y(t) = \lambda \cdot e^{\ln(t)}$ qui revient à $y(t) = \lambda t$.

Question 2 : BCDE

On considère le modèle épidémiologique SIR, en prenant en compte le taux de contamination (k), le taux de retrait (r) et la mortalité (m) qui sont des constantes.

On prend également en compte un flux entrant de population « saine », qui lui est proportionnel au nombre de Susceptibles.



Le système sans vaccination est le suivant :

$$\begin{cases} \frac{dS}{dt} = f.S - k.SI - m.S \\ \frac{dI}{dt} = k.SI - r.I - m.I \\ \frac{dR}{dt} = r.I - m.R \end{cases}$$

A FAUX Tout est vrai sauf que le système est sans second membre car le flux entrant est proportionnel à S.

B VRAI Effectivement, la maladie est contagieuse, ce qui met en dépendance S et I car si S=0 ou I=0, on ne peut pas avoir de propagation de la maladie. Il faut au moins 1 malade (I) pour propager la maladie, mais également des personnes susceptibles de devenir malade (S) pour qu'il y ait une propagation et donc une épidémie.

C VRAI Voir cours : les retirés (R) sont des personnes mortes, mises en quarantaine ou immunisées.

D VRAI Le système devient alors :

$$\begin{cases} \frac{dS}{dt} = f.S - k.SI - m.S - v.S \\ \frac{dI}{dt} = k.SI - r.I - m.I \\ \frac{dR}{dt} = r.I - m.R + v.S \end{cases}$$

E VRAI Les personnes vaccinées seront immunisées, ainsi comme on l'a vu à l'item C, cela correspond au retirés. On peut également le voir par l'équation différentielle $\frac{dR}{dt}$.

Question 3 : ABCE

On note les évènements de la façon suivante :

- W1 : « utiliser le réseau d'eau W1 »
- W2 : « utiliser le réseau d'eau W2 »
- W3 : « utiliser le réseau d'eau W3 »
- C : « avoir le choléra »

Les données de l'énoncé sont :

$$P(W1) = 0,3 \quad P(W2) = 0,15 \quad P(W3) = 0,55 \quad P_{W1}(C) = \frac{1}{6} \quad P_{W2}(C) = \frac{1}{3} \quad P_{W3}(C) = \frac{2}{11}$$

A VRAI On cherche la probabilité d'avoir le choléra donc P(C) :

$$P(C) = P(C \cap W1) + P(C \cap W2) + P(C \cap W3) \quad (\text{formule des probabilités totales})$$

$$P(C) = P_{W1}(C) \times P(W1) + P_{W2}(C) \times P(W2) + P_{W3}(C) \times P(W3) \quad (\text{formule des probabilités composées})$$

$$P(C) = \frac{1}{6} \times 0,3 + \frac{1}{3} \times 0,15 + \frac{2}{11} \times 0,55$$

$$P(C) = \frac{0,3}{6} + \frac{0,15}{3} + \frac{1,1}{11} = 0,05 + 0,05 + 0,1 = 0,2$$

B VRAI On cherche la probabilité pour que deux individus indépendants utilisent le même réseau d'eau donc soit qu'ils utilisent tous les deux W1 soit W2 soit W3. De plus, ils sont indépendants donc $P(W1 \cap W1) = P(W1) \times P(W1)$.

On a donc :

$$\begin{aligned} P(W1 \cap W1) + P(W2 \cap W2) + P(W3 \cap W3) &= P(W1) \times P(W1) + P(W2) \times P(W2) + P(W3) \times P(W3) \\ &= 0,3 \times 0,3 + 0,15 \times 0,15 + 0,55 \times 0,55 \\ &= 0,09 + 0,0225 + 0,3025 \\ &= 0,415 \end{aligned}$$

Et $0,40 < 0,415 < 0,42$

C VRAI Soient 4 habitants indépendants les uns des autres : 1, 2, 3 et 4. On cherche la probabilité que 2 d'entre eux soient atteints du choléra, soit l'évènement suivant : « $C \cap C \cap \bar{C} \cap \bar{C}$ ».

$$P(C \cap C \cap \bar{C} \cap \bar{C}) = P(C) \times P(C) \times P(\bar{C}) \times P(\bar{C}) = 0,2 \times 0,2 \times (1-0,2) \times (1-0,2) = 0,2 \times 0,2 \times 0,8 \times 0,8 = 0,0256$$

Or il existe 6 répartitions possibles pour les 4 habitants :

- $C1 \cap C2 \cap \bar{C3} \cap \bar{C4}$
- $C1 \cap C3 \cap \bar{C2} \cap \bar{C4}$
- $C1 \cap C4 \cap \bar{C2} \cap \bar{C3}$
- $C2 \cap C3 \cap \bar{C1} \cap \bar{C4}$
- $C2 \cap C4 \cap \bar{C1} \cap \bar{C3}$
- $C3 \cap C4 \cap \bar{C1} \cap \bar{C2}$

Donc la probabilité d'avoir 2 habitants atteints du choléra parmi ces 4 habitants est de $6 \times 0,0256$.

D FAUX On veut savoir si les évènements C et W1 sont indépendants. D'après le cours, ils sont indépendants si $P(C \cap W1) = P(C) \times P(W1)$.

$$P(C \cap W1) = 0,05 \text{ (on l'a calculé dans l'item A)}$$

$$P(C) \times P(W1) = 0,2 \times 0,3 = 0,06$$

$$\text{Or } 0,05 \neq 0,06$$

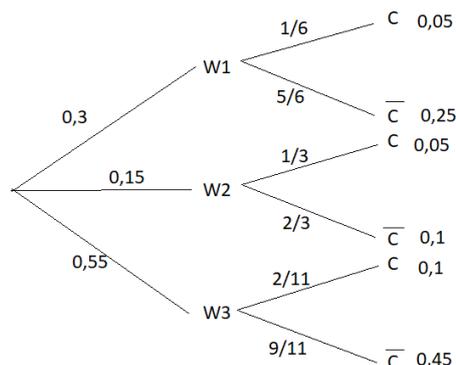
$$\text{Donc } P(C \cap W1) \neq P(C) \times P(W1)$$

Donc C et W1 ne sont pas indépendants.

E VRAI On cherche $P_C(W1)$.

$$P_C(W1) = \frac{P(C \cap W1)}{P(C)} = \frac{0,05}{0,2} = \frac{1}{4}$$

Il est aussi possible de résoudre cet exercice en utilisant un arbre de probabilités :



Question 4 : AC

On a : $W = aX + bY$, or d'après les formules du cours : $W (a\mu_X + b\mu_Y ; \sqrt{a^2\sigma_X^2 + b^2\sigma_Y^2})$

On calcule maintenant l'espérance et l'écart-type de W en considérant $a = -1$ et $b = 2$.

$$\mu_W = -20 + 2 \times 10 = 0$$

$$\sigma_W = \sqrt{3^2 + 2^2 \cdot 2^2} = \sqrt{9 + 16} = \sqrt{25} = 5$$

A VRAI W a pour espérance 0, ainsi c'est une variable aléatoire centrée. Cependant elle n'est pas réduite car son écart-type est différent de 1.

B FAUX Voir le calcul plus haut.

C VRAI On va centrer réduire la variable aléatoire Y pour trouver la valeur de notre probabilité dans la table : fonction de répartition de la loi normale centrée réduite.

$$P(Y \geq 8) = P\left(\frac{Y-10}{2} \geq \frac{8-10}{2}\right) = P(Z \geq -1) = P(Z \leq 1) = 0,8413$$

On a cherché notre valeur dans la table pour un $z = 1$, on trouve donc 0,8413.

D FAUX On peut réfuter cet item si on arrive à visualiser la courbe de Gauss, en s'intéressant à l'aire sous la courbe :

$$P(15 < X < 20) = P(X < 20) - P(X < 15) = P(X < 20) - (1 - P(X > 15)) = P(X < 20) - 1 + P(X > 15)$$

E FAUX L'intervalle de fluctuation au risque de 5% de W revient à l'intervalle contenant 95% des valeurs de W . Donc $[\mu_W - 2\sigma_W ; \mu_W + 2\sigma_W] = [0 - 2 \times 5 ; 0 + 2 \times 5] = [-10 ; 10]$.

Question 5 : AE

Dans un exercice sur les intervalles, il faut d'abord identifier la situation et le type d'intervalle recherché.

Ici, on dispose d'un échantillon à partir duquel on veut déterminer les paramètres de la population. On recherche donc un intervalle de confiance.

Il faut vérifier que les conditions d'application pour un intervalle de confiance sont validées, c'est-à-dire que $n \geq 30$.

Pour chaque vaccin, $n = 100 \geq 30$ donc on peut calculer notre intervalle de confiance.

A VRAI La formule de l'intervalle de confiance est :

$$ic_{1-\alpha}(p) = f \pm z_{\alpha/2} \times \sqrt{\frac{f(1-f)}{n}}$$

$$ic_{0,95}(p_{vdd}) = \frac{85}{100} \pm 2 \times \sqrt{\frac{0,85 \times 0,15}{100}}$$

$$ic_{0,95}(p_{vdd}) = 0,85 \pm \frac{2 \times 0,357}{10} = 0,85 \pm \frac{0,714}{10} = 0,85 \pm 0,0714$$

$$ic_{0,95}(p_{vdd}) = [0,7786; 0,9214]$$

Pour les intervalles, on minore la valeur inférieure et on majore la valeur supérieure pour être sûr qu'il contienne tous les résultats possibles donc :

$$ic_{0,95}(p_{vdd}) = [0,77; 0,93]$$

Il vaut vérifier les conditions à posteriori :

- $nf_1 = 100 \times 0,77 = 77 \geq 5$
- $n(1 - f_1) = 100 \times 0,23 = 23 \geq 5$
- $nf_2 = 100 \times 0,93 = 93 \geq 5$
- $n(1 - f_2) = 100 \times 0,07 = 7 \geq 5$

Elles sont respectées donc l'intervalle est valable.

B FAUX L'intervalle donné est faux car on n'a pas arrondi correctement les valeurs (voir correction item A).

C FAUX On ne recherche pas un intervalle de fluctuation mais un intervalle de confiance.

D FAUX On fait un intervalle de confiance pour trouver la proportion théorique dans la population p à partir de la proportion estimée dans l'échantillon f . On cherche donc $ic_{0,95}(p_{vdd})$ et pas $ic_{0,95}(f_{vdd})$.

E VRAI Plus le niveau de confiance est élevé, plus l'intervalle de confiance est large. En effet, un intervalle de confiance à 99% va englober plus de valeurs qu'un intervalle de confiance à 95% pour avoir moins de chances de se tromper. Donc un intervalle de confiance à 99% est plus large qu'un intervalle de confiance à 95%.

Question 6 : BDE

Pour cet exercice de test statistique, il faut tout d'abord **poser l'hypothèse nulle H_0** : il n'y a pas de différence entre les deux proportions : $p_{vdd} = p_{vnd}$

Ainsi que l'hypothèse alternative H_1 qui est dans notre cas **bilatéral, c'est à dire qu'on teste juste si les proportions sont significativement différentes : $p_{vdd} \neq p_{vnd}$ sans s'intéresser au sens de la différence.**

D'après ces données on peut dire la réponse **B est vraie**. Cependant la **réponse A est fausse** car on ne teste pas notre hypothèse avec les estimations des proportions calculées notées f mais avec les proportions théoriques de succès dans les populations notées p .

C FAUX : Les conditions d'application du test de comparaison de deux proportions observées sont :

$$\begin{aligned} n_{vdd}p_0 &\geq 5 & n_{vdd}(1 - p_0) &\geq 5 \\ n_{vnd}p_0 &\geq 5 & n_{vnd}(1 - p_0) &\geq 5 \end{aligned}$$

Avec $p_0 = \frac{n_{vdd}f_{vdd} + n_{vnd}f_{vnd}}{n_{vdd} + n_{vnd}}$ on comprend donc que p_0 est différent de f_{vdd} et de f_{vnd} .

On vérifie à présent que ces conditions sont validées :

$$f_{vdd} = 0,85 ; f_{vnd} = 0,9 ; n_{vdd} = 100 ; n_{vnd} = 100$$

$$p_0 = \frac{100 \times 0,85 + 100 \times 0,9}{100 + 100} = \frac{85 + 90}{200} = \frac{175}{200} = 0,875$$

On a alors $0,875 \times 100 = 87,5 \geq 5$ et $1 - 0,875 \times 100 = 0,125 \times 100 = 12,5 \geq 5$

Les conditions sont donc bien vérifiées.

On calcule à présent le degré de significativité du test p (sachant que ce dernier est bilatéral)

$$p = P(|Z| > |z|) = 2 \times P(|Z| > |z|) = 2 \times (1 - \varphi(|z|)) = 2 \times (1 - \varphi(1,07)) = 2 \times (1 - 0.8577) \approx 0,28$$

On lit la valeur p à partir de la statistique de test dans la table de répartition de la loi normale centrée réduite par interpolation linéaire. **Donc D VRAI**

E VRAI car $0,28 \geq 0,1$ donc $p \geq \alpha$ donc on ne peut pas rejeter l'hypothèse nulle, on conclut alors qu'il n'y a pas de différence significative entre les 2 traitements au risque $\alpha = 10\%$.

NB : lorsqu'on prend le risque α , on le compare à p, on peut alors se dire que la différence est significative quand le risque n'est pas dépassé : **$p \leq \alpha$: rejet H_0**

A contrario quand on utilise le seuil on peut dire que la différence est significative à partir du moment où le seuil est dépassé : **$|Z_{calculé}| > |Z_{seuil}|$: rejet de H_0**

Question 7 : AD

Les données de l'énoncé sont :

$$n = 106 \quad n_B = 76 \quad n_C = 30 \quad m_B = 50 \quad m_C = 85 \quad s_B = 38 \quad s_C = 30$$

Les hypothèses à tester sont :

$$H_0 : \mu_B = \mu_C \text{ et } H_1 : \mu_B \neq \mu_C$$

On est donc dans la situation d'un test bilatéral qui doit comparer 2 moyennes dans 2 échantillons.

A VRAI $n_B = 76 \geq 30$ et $n_C = 30 \geq 30$ donc on peut calculer la statistique de test suivante qui suit la loi normale centrée réduite $N(0 ; 1)$:

$$z = \frac{m_C - m_B}{\sqrt{\frac{s_C^2}{n_C} + \frac{s_B^2}{n_B}}} = \frac{85 - 50}{\sqrt{\frac{30^2}{30} + \frac{38^2}{76}}} = \frac{35}{\sqrt{30 + \frac{38}{2}}} = \frac{35}{\sqrt{30 + 19}} = \frac{35}{\sqrt{49}} = \frac{35}{7} = 5$$

B FAUX Voir correction de l'item A

C FAUX On est dans le cas d'un test bilatéral donc :

$$p = P(|Z| > z) = P(|Z| > 5) = 2 \times (1 - \varphi(5))$$

Dans la table de répartition de la loi normale centrée réduite, la plus grande valeur de z possible est 4,09 or on a $z = 5$. On va utiliser la valeur correspondant à 4,09, en sachant que la valeur correspondant à 5 sera plus grande.

3,7	0,999907	0,999900	0,999890	0,999870	0,999851	0,999831	0,999812	0,999792	0,999772	0,999752
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

Donc $\varphi(5) \approx 0,99998$ donc :

$$p = 2 \times (1 - 0,99998) = 2 \times 0,00002 = 0,00004$$

$$p = 0,00004 < 0,05 = \alpha$$

Donc on rejette H_0 et on peut conclure que la différence entre les paramètres moyens d'élasticité est statistiquement significative au risque $\alpha = 5\%$.

D VRAI On a conclu que la différence entre les paramètres moyens est statistiquement significative au risque $\alpha = 5\%$. En l'occurrence, le paramètre moyen d'élasticité des nodules cancéreux est supérieur au paramètre moyen d'élasticité des nodules bénins, ce qui nous informe sur le sens de la différence

entre les 2 paramètres : le paramètre d'élasticité moyen des nodules cancéreux est significativement plus élevé que celui des nodules bénins au risque alpha de 5%.

E FAUX voir correction item C et D.

Question 8 : BCD

On pose : Patient malade M : quand le nodule est cancéreux, sinon non malade NM.

De plus, on est devant un test effectué qui permet de classer les nodules selon une échelle de 1 à 5, sachant qu'il s'avère positif pour un seuil supérieur ou égal à 4 (cancer ou suspicion de cancer).

On sait qu'on a au total 106 patients adressés pour une ablation chirurgicale de leur nodule et qu'on a au total 43 tests positifs. De plus $Sp = 75\% = \frac{57}{76} = \frac{VN}{VN+FP}$ la spécificité correspond à la probabilité qu'un sujet soit négatif au test sachant qu'il n'est pas réellement malade.

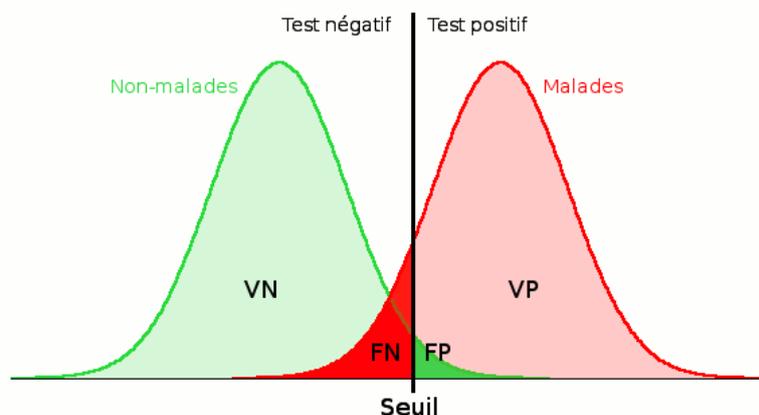
A partir de ces données on peut remplir le **tableau de contingence** en faisant des sommes/différences (résultats des tests diagnostiques très souvent présentés sous cette forme !!)

	Malades M	Non Malades NM ou M	Totaux
T +	24	19	43
T -	6	57	63
Totaux	30	76	106

Ainsi **Sensibilité** = $\frac{VP}{VP+FN} = \frac{24}{30} = 0,8$ donc **A FAUX** et **B VRAI**

C VRAI En effet si le seuil de positivité du test est plus bas, on aura alors plus de patients ayant un test positif (par exemple seuil ≥ 3 , on aura les patients ayant un score de 3, de 4 et de 5), le nombre de VP augmente alors et donc la Sensibilité aussi.

On voit bien tout cela sur le graphique ci-dessous :



$$\begin{aligned}
\text{D VRAI Proba post test} &= \frac{\text{Odds post test}}{1 + \text{Odds post test}} = \frac{\text{Odds pré test} \times RV+}{1 + \text{Odds pré test} \times RV+} \\
&= \frac{P(M) / P(\bar{M}) \times P(T^+ | M) / P(T^+ | \bar{M})}{1 + [P(M) / P(\bar{M}) \times P(T^+ | M) / P(T^+ | \bar{M})]} \\
&= \frac{(30/106) / (76/106) \times (24/30) / (19/76)}{1 + [(30/106) / (76/106) \times (24/30) / (19/76)]} \\
&= \frac{30/76 \times 0,8/0,25}{1 + [30/76 \times 0,8/0,25]} \approx 0,56
\end{aligned}$$

On pouvait aussi trouver cette probabilité grâce à un deuxième calcul (cependant le "1+" de l'énoncé nous indiquait la formule à utiliser pour obtenir 0,56)

$$\text{Ainsi Probabilité post test qd test+} = \text{VPP} = P(M | T^+) = \frac{P(T^+ | M) \times P(M)}{P(T^+ | M) \times P(M) + P(T^+ | \bar{M}) \times P(\bar{M})}$$

(formule de Bayes donnée dans le formulaire)

En remplaçant les termes de cette formule par les probabilités de notre tableau on retrouvait 0,56 cependant en passant par des calculs beaucoup plus longs qu'avec la formule précédente vu qu'on retrouvait les chiffres de l'énoncé.

E FAUX : Dans cet exercice il était dit que l'échantillon correspondait à 106 patients adressés pour ablation d'un nodule thyroïdien. La répartition malades /non malades est le reflet de la prévalence de la maladie dans cette population. Ainsi notre **échantillon est représentatif de la maladie, on peut donc estimer les probabilités post test, VPP et VPN**. Avec un échantillon non représentatif, comme dans le cas d'une étude cas/témoin (constitution d'un groupe de malades et d'un groupe de sujets non malades indépendamment) ça n'aurait pas eu de sens d'estimer ces valeurs.

Question 9 : ADE

A VRAI C'est une étude de cohorte car il est dit clairement que les sujets ont été suivis durant un certain temps (5 ans) et dans l'attente de l'apparition d'un évènement (ici la récurrence de leur cancer).

B FAUX voir A. Une étude transversale est une photographie de la population à un moment donné, sans suivi des sujets.

C FAUX On pose r = patients qui ont eu une récurrence et \bar{r} = patients qui n'ont pas eu une récurrence. Le taux de récurrence, correspond au taux d'incidence de la récurrence :

$$\frac{r}{r \times \frac{t}{2} + \bar{r} \times t} = \frac{60}{60 \times 2,5 + 240 \times 5} = \frac{60}{1350} \text{ personnes années.}$$

D VRAI Le risque de récurrence à 5 ans correspond au rapport entre les patients qui ont eu une récurrence et la totalité des patients de l'étude : $\frac{r}{n} = \frac{60}{300} = 0,2$.

E VRAI L'odds de récurrence à 5 ans correspond au rapport entre les patients qui ont eu une récurrence et les patients qui n'en ont pas eu : $\frac{r}{\bar{r}} = \frac{60}{240} = 0,25$.

Question 10 : AC

Tout d'abord on relève les informations importantes de l'énoncé : On est devant un score de Gleason qui mesure l'agressivité de la tumeur chez 300 patients inclus dans une étude.

Parmi les 100 patients ayant un score ≥ 8 , on compte 24 récurrences au bout de 5 ans de suivi.

Parmi les 200 patients ayant un score < 8 , on compte 36 récurrences au bout de 5 ans de suivi.

De plus la statistique de test du Chi-2 est : $\chi^2 = 1,5$

A VRAI Il y a un suivi des patients donc c'est bien une cohorte, de plus on s'intéresse à la force d'association d'un facteur ici de pronostic (score de Gleason) à la maladie, donc l'étude est bien analytique.

B FAUX On a $\chi^2 = 1,5$, pour savoir si la différence de risque de récurrence est significativement différente entre les deux groupes, on cherche le degré de significativité p dans la table du chi-2, à $(c-1)(l-1)$ ddl soit $(2-1)(2-1)$ ddl donc à **1ddl**.

La table du chi-2 nous donne les valeurs pour un test bilatéral ce qui est bien notre cas ici donc pas de modifications à faire. Si le test avait été unilatéral on aurait divisé p par 2 ou multiplié α par 2.

On lit donc 1,5 entre 0,750 et 0,9 donc : $(1-0,750) < p < (1-0,9)$ soit **0,25 > p > 0,1**

(On fait 1 moins le degré de significativité lu dans la table car cette dernière nous donne p tel que $p = P(X < k)$ or nous cherchons p tel que $p = P(\chi^2 > k)$.)

$\alpha = 0,05$ soit $\alpha \leq p$ donc on ne peut pas rejeter l'hypothèse nulle d'absence de différence. La différence n'est donc pas significative au risque alpha bilatéral de 5%.

C VRAI

Risque Relatif = $\frac{\text{Risque de récurrence chez les patients ac Score } \geq 8}{\text{Risque de récurrence chez les patients ac Score } < 8}$

$$= \frac{24/100}{36/200} = \frac{24}{100} \times \frac{200}{36} = \frac{8 \times 3 \times 2}{4 \times 9} = \frac{4}{3}$$

D FAUX : le risque relatif est bien supérieur à 1, cependant il n'est **pas significativement supérieur à 1** car le test du chi-2 ne permettait pas de conclure à une différence significative.

E FAUX : Pour les mêmes raisons que précédemment on ne peut pas conclure que le score de Gleason est un facteur pronostique de récurrence : absence de différence significative avec le test effectué.

(Cependant cela n'exclut pas qu'avec un autre test ou un autre risque α les résultats n'auraient pas été significatifs.)

Question 11 : ACD

A VRAI $p_{X,Y} = \frac{s_{X,Y}}{s_X \cdot s_Y} = \frac{27}{6 \times 30} = \frac{3 \times 9}{6 \times 3 \times 10} = \frac{9}{6 \times 10} = \frac{0,9}{6} = 0,15$

B FAUX $b_1 = \frac{s_{X,Y}}{s_X \cdot s_X} = \frac{27}{30^2} = \frac{0,9}{30} = 0,03$

C VRAI Les tests des coefficients de corrélation et de régression sont testés à $n-2$ ddl. Or $n=227$, ainsi on teste à 225 ddl.

D VRAI Il faut calculer la statistique de test :

$$t = \frac{b_1}{\sqrt{\frac{s_y^2 - b_1^2}{s_x^2 - b_1^2} \frac{1}{n-2}}} = \frac{0,03}{\sqrt{\frac{36 - 0,0009}{900} \frac{1}{225}}} = \frac{0,03}{\sqrt{\frac{0,04 - 0,0009}{225}}} = \frac{0,03}{\sqrt{\frac{0,039}{225}}} = \frac{0,03}{\sqrt{\frac{0,04}{225}}} = \frac{0,03}{\frac{0,2}{15}} = \frac{0,03 \times 15}{0,2} = \frac{0,3 \times 15}{2} = 2,25$$

Il faut maintenant rechercher notre valeur dans la table de la loi de Student, à 225 ddl, or elle s'arrête à 100, on se reporte donc à la ligne ∞ .

On cherche le t_{seuil} pour $\alpha = 0,05$, on trouve 1,96. Or notre valeur test est de 2,25, elle est donc supérieure au seuil. Ainsi on rejette H_0 , le coefficient de régression est donc statistiquement significatif au risque alpha bilatéral de 5%.

E FAUX X explique $p_{X,Y}^2$ de Y. Il faut faire attention à ne pas oublier le carré. C'est un piège récurrent. Ainsi la variabilité du taux de filtration glomérulaire explique 2,25% de la variabilité de la concentration de vitamine D sérique.

Question 12 : ACD

On a un test bilatéral avec $p_G = 0,3$ et $p_D = 0,4$.

A VRAI On utilise la formule suivante :

$$n = \frac{(z_{1-\beta} \sqrt{p_G(1-p_G)} + p_D(1-p_D) + z_{1-\alpha/2} \sqrt{2p_0(1-p_0)})^2}{\delta^2}$$

p_0 est la proportion d'échec attendue sous l'hypothèse H_0 de non-différence entre les 2 traitements soit : $p_0 = \frac{p_G + p_D}{2} = \frac{0,3 + 0,4}{2} = 0,35$.

δ correspond à l'écart entre les deux proportions soit : $\delta = p_D - p_G = 0,1$

Par lecture dans les tables : $z_{1-\beta} \approx 1$ et $z_{1-\alpha/2} \approx 2$

Donc :

$$n = \frac{(1\sqrt{0,3 \times 0,7} + 0,4 \times 0,6 + 2\sqrt{2 \times 0,35 \times 0,65})^2}{0,1^2} = \frac{(\sqrt{0,45} + 2\sqrt{0,45})^2}{0,01} = \frac{(3\sqrt{0,45})^2}{0,01} = \frac{9 \times 0,45}{0,01} = \frac{4,05}{0,01} = 405$$

On a donc besoin de 405 patients par bras (et non pas 405 patients au total).

B FAUX Le critère de jugement principal est le décès ou la progression du sarcome à 24 mois donc un patient en échec est soit décédé soit a connu une progression de son sarcome à 24 mois.

C VRAI La puissance correspond à $1 - \beta$. C'est la probabilité de rejeter H_0 et que H_0 soit fausse.

Dans l'énoncé, cela correspond à « on souhaite pouvoir conclure à une différence statistiquement significative entre les 2 bras avec une probabilité de 85% » donc $1 - \beta = 0,85$.

D VRAI Le critère de jugement principal est le décès ou la progression à 24 mois donc un patient en succès n'est ni décédé ni n'a eu une progression de son sarcome. Autrement dit, un patient en succès est vivant sans progression.

E FAUX $1 - \beta = 0,85$ donc $\beta = 0,15$. Ici, c'est le risque de première espèce qui vaut 5%, soit α , c'est-à-dire le risque de rejeter H_0 et que H_0 soit vraie.

Question 13 : AD

Tout d'abord on fixe les paramètres de l'exercice :

La mortalité associée à 2 différents traitements est estimée par un modèle exponentiel à taux proportionnels.

λ est le taux de mortalité, **Dxb : doxorubicine**, et **Gdt : gemcitabine + docetaxel**

$\lambda_{Dxb} = 2,4 \text{ an}^{-1}$ et $\lambda_{Gdt} = \frac{1}{3} \times \lambda_{Dxb}$ car le taux relatif de mortalité est égal à $\frac{1}{3}$ avec un meilleur pronostic pour le traitement Gdt donc ce dernier a un taux de mortalité 3 fois moins élevé.

A partir de ces données on a aussi $S_{Gdt} = S_{Dxb}^{1/3}$ (S signifiant Survie)

Rappel de la formule : $S_0(t) = e^{-\lambda t}$

A VRAI $S_{Dxb}(6 \text{ mois}) = S_{Dxb}(0,5 \text{ an}) = e^{-0,5 \times 2,4} = (e^{-2,4})^{0,5} = 0,09^{0,5} = 0,3$
(Mettre à la puissance 0,5 revient à faire la racine carrée)

NB : il y a une petite erreur dans les aides au calcul donnés dans l'énoncé, en effet $\ln(0,09) \neq 2,4$ c'est plutôt **$\ln(0,09) \approx -2,4$** (oubli du signe moins...)

B FAUX

$$S_{Gdt}(18 \text{ mois}) = S_{Gdt}(1,5 \text{ ans}) = (e^{-1,5 \times 2,4})^{1/3} = e^{\frac{3}{2} \times 2,4 \times \frac{1}{3}} = e^{-2,4 \times \frac{1}{3}} = (e^{-2,4})^{1/2} = 0,3$$

C FAUX Simple question de cours, dans les modèles exponentiels à taux proportionnels, ce sont les taux de mortalité qui sont proportionnels (donc avec le rapport des taux égal à $\frac{1}{3}$) et non pas les risques/probabilités.

D VRAI $S_{Dgt}(1 \text{ an}) = e^{-2,4 \times 1 \times 1/3} = e^{-0,8} \approx 0,44$

E FAUX C'est l'inverse, la survie estimée à 1 an dans le bras Doxorubicine est égale à celle du bras Gemcitabine + Docetaxel au cube.

$$S_{Gdt} = S_{Dxb}^{1/3} \text{ soit } S_{Dxb} = S_{Gdt}^3$$

Question 14 : ANNULE