

BTSA 2015 : corrigé de traitement de données, juin

Proposition de corrigé - non officielle. BTSA, toutes options, traitement de données, session 2015, Métropole, Antilles, Guyane et Réunion, code 2015-BTS103-NOR-ME-AN-GU-RE. Les trois exercices sont traités.

[Sujet officiel](#) · [Annales de ce concours ou examen](#)

Exercice 1 - Fruits et légumes

1. Toutes les cases du tableau sont positives ou nulles. Leur somme est **1** : il s'agit bien d'une loi conjointe de probabilité.

2. Ne consommer ni fruit ni légume correspond à $X = 0$ et $Y = 0$. La case correspondante vaut **0**.

3 et 5. On additionne les colonnes pour la loi de X et les lignes pour celle de Y :

Valeur	0	1	2	3
$P(X = \text{valeur})$	0,11	0,36	0,36	0,17
$P(Y = \text{valeur})$	0,09	0,32	0,30	0,29

Ainsi $P(X = 0) = 0,11$ et $P(Y = 0) = 0,09$. Chaque ligne de cette nouvelle table somme à 1.

4. Si X et Y étaient indépendantes, $P(X = 0 \text{ et } Y = 0)$ vaudrait $0,11 \times 0,09 = 0,0099$. Or cette probabilité vaut 0 dans le tableau. Les deux variables **ne sont pas indépendantes**.

6a et 6b. $Z = X + Y$ compte le nombre total de variétés de fruits et légumes consommées la veille. Pour une valeur de Z , on additionne les cases dont les indices ont cette somme : par exemple $P(Z = 2) = 0,03 + 0,12 + 0,05 = 0,20$.

z	0	1	2	3	4	5	6
$P(Z = z)$	0	0,03	0,20	0,32	0,28	0,15	0,02

6c. $P(Z \geq 5) = 0,15 + 0,02 = 0,17$. Dans cette enquête, 17 % des personnes ont consommé au moins cinq variétés au total.

7. $E(Z) = 1 \times 0,03 + 2 \times 0,20 + 3 \times 0,32 + 4 \times 0,28 + 5 \times 0,15 + 6 \times 0,02 = 3,38$. Cela représente une moyenne de 3,38 variétés par personne, et non une quantité entière consommée par chaque personne.

Exercice 2 - Proportion de personnes dépassant le seuil du sujet

Partie A. Estimation sur 1 000 personnes

1a. F est la moyenne de 1 000 indicatrices de succès : $E(F) = p$.

1b. Sous les hypothèses de l'énoncé, la loi de F est approximativement normale, de moyenne p et de variance $p(1 - p)/1\,000$, donc d'écart type $\sqrt{p(1 - p)/1\,000}$. L'approximation convient lorsque les nombres attendus de succès et d'échecs sont suffisamment grands.

2. La fréquence observée est $f = 0,168$. L'intervalle de confiance usuel à 95 % est $f \pm 1,96\sqrt{f(1 - f)/1\,000}$. On trouve $[0,1448275 ; 0,1911725]$, soit, au millièm, **[0,145 ; 0,191]**. La proportion est estimée entre environ 14,5 % et 19,1 %.

Partie B. Probabilité fixée à 0,17

1. La probabilité d'un échec est 0,83. Pour dix prélèvements indépendants, la probabilité de dix échecs vaut $0,83^{10} \approx 0,15516$, soit **0,16** au centièm.

2a. Le nombre de succès parmi 200 individus suit **$B(200 ; 0,17)$** .

2b. L'espérance vaut **34**, la variance **28,22** et l'écart type $\sqrt{28,22} \approx 5,31225$. Les effectifs attendus 34 et 166 sont grands : l'approximation normale est adaptée.

2c. Pour inclure les valeurs entières de 30 à 40, la correction de continuité donne **$P(29,5 \leq Y \leq 40,5)$** . Après centrage-réduction, la probabilité est $\Phi((40,5 - 34)/5,31225) - \Phi((29,5 - 34)/5,31225) \approx 0,69098$, soit **0,69** au centièm. La somme binomiale exacte, environ 0,68772, donne le même arrondi.

Exercice 3 - Consommation de pain

1. Ajustement direct

Le coefficient de corrélation de T et M vaut $r \approx -0,933$. La liaison est décroissante, mais le graphique montre une diminution qui ralentit : les points suivent une courbe. La droite passe sous les derniers points et conduirait ensuite à des masses négatives. Cet ajustement direct convient mal.

2. Ajustement après logarithme

2a. Pour $Z = \ln(M)$, on trouve **$r(T, Z) \approx -0,982$** . Sa valeur absolue est plus proche de 1, ce qui indique un meilleur alignement après cette transformation.

2b. La régression de Z en T donne **$\hat{z} = -0,0181120611t + 6,7099205686$** . Les coefficients demandés au millièm s'écrivent -0,018 et 6,710, mais on conserve davantage de chiffres pour la prévision.

2c. En prenant l'exponentielle, **$m \approx \exp(6,7099205686 - 0,0181120611t)$** . Ce modèle conserve une masse positive et traduit un ralentissement de la baisse en grammes.

2d. En 1935, $t = 1935 - 1900 = 35$. Avec les coefficients non arrondis, on obtient **$m \approx 435,284$ g par jour et par adulte**. Il s'agit d'une estimation interpolée entre les observations, pas d'une valeur historique directement relevée.