

# Probabilités et statistiques pour l'ingénieur

## M1 Informatique

Examen, janvier 2018

Ce devoir surveillé consiste en **?? exercices** sur ?? pages.

Les calculatrices ne sont pas autorisées.

Les exercices peuvent être traités dans le désordre.

Penser à bien **justifier** les réponses. La notation prendra en compte le **soin et la clarté des réponses**.

### Exercice 1.

Trois machines A, B, et C produisent respectivement 50%, 30% et 20% du nombre total de pièces produites dans une usine. Le pourcentage de pièces défectueuses de ces machines est respectivement de 3%, 4% et 5%.

1. Une pièce aléatoire est extraite : déterminez la probabilité qu'elle soit défectueuse.
2. Si une pièce est défectueuse, quelle est la probabilité que cette pièce ait été produite par la machine A.
3. L'usine produit 1000 pièces par jour. Chaque pièce défectueuse est éliminée, ce qui entraîne une perte de €0.5 par pièce. Quel est la perte économique attendue au bout d'une journée de production ?

### Exercice 2.

Soit  $X$  une variable aléatoire continue avec densité

$$f(x) = \begin{cases} x & \text{si } 0 \leq x \leq \sqrt{2} \\ 0 & \text{ailleurs} \end{cases}$$

1. Tracer le graphique de  $f$ .
2. Calculer  $\mathbb{P}(X \leq 1)$ .

### Exercice 3.

On considère  $n$  réalisations indépendantes et identiquement distribuées  $X_1, \dots, X_n$  d'une variable aléatoire  $X$  d'espérance  $E[X] = \mu$  et variance  $V(X) = \sigma^2$ . Soit  $\bar{X} = \frac{X_1 + \dots + X_n}{n}$  la moyenne.

1. Montrer que  $E[\bar{X}] = \mu$  à l'aide des propriétés de l'espérance.
2. Montrer que  $V(\bar{X}) = \frac{\sigma^2}{n}$  à l'aide des propriétés de la variance.
3. Quelle est la loi de  $Z = \frac{\bar{X} - \mu}{\sigma} \sqrt{n}$  quand  $n$  est grand ? Justifier en citant un important résultat théorique vu en cours.
4. On sait que  $\mathbb{P}(-1.96 \leq Z \leq 1.96) = 0.95$ . À l'aide de cela trouver l'expression de l'intervalle de confiance à 95% pour  $\mu$ , c'est à dire trouver deux variables aléatoires  $L$  et  $U$  t.q.  $\mathbb{P}(L \leq \mu \leq U) = 0.95$ .
5. On suppose  $\sigma = 20$ . Donner le plus petit  $n$  t.q. la longueur de l'intervalle soit inférieur à 10. Indication : pour faciliter le calcul, on pourra prendre  $1.96 \simeq 2$ .

### Exercice 4.

On s'intéresse au montant  $X$  des dépenses en biens et services culturels (livres, spectacles, musées...) au sein d'une population. Dans un premier temps, on cherche à savoir si le montant moyen  $\mu$  est inférieur à un seuil fixé à €1500. Pour cela on fait une enquête et on collecte les observations i.i.d.  $x_1, \dots, x_{1000}$ . On décide ensuite de réaliser un test statistique.

1. Donner les hypothèses  $H_0$  et  $H_1$ .
2. Donner le nom du test et sa latéralité. A-t-on besoin de supposer  $X$  gaussienne pour ce test ?
3. Donner la commande R pour réaliser le test.
4. La probabilité critique associée au test est 0.02. Conclure.

Après avoir réalisée cette étude, le gouvernement intervient avec une campagne pour inciter la consommation culturelle. On s'intéresse à savoir si cette campagne est efficace et on recontacte les 1000 individus précédents pour collecter leur nouvelles dépenses. Ensuite on fait un nouveau test pour savoir si le montant moyen a augmenté.

5. Donner les hypothèse de ce deuxième test.
6. Donner le nom du test et sa latéralité. Dire si les échantillons sont appariés (justifier sa réponse).

On veut enfin savoir si il y a une dépendance entre  $X$  est le revenu moyen  $Z$ .

7. Donner les hypothèses et le nom de ce troisième test.

### Exercice 5.

On s'intéresse aux facteurs censés affecter la vitesse des voitures.

1. Dans un premier temps, on s'intéresse à la relation entre le poids des voitures  $x$  en tonnes et leur vitesse  $y$  exprimée par le temps en secondes nécessaires à un véhicule arrêté pour atteindre la distance de 400 mètres. Pour cela on a relevé 32 observations  $(x, y)$  : voir Fig. 1, plot de gauche.

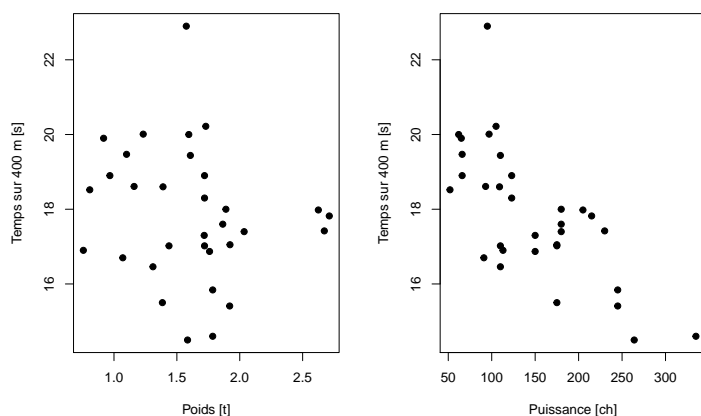


FIGURE 1 – Plot à gauche : poids  $x$  et vitesse  $y$ . Plot à droite : puissance  $x'$  et vitesse  $y$ .

On modélise la relation entre  $x$  et  $y$  à l'aide d'un modèle de régression linéaire.

- (a) Ecrire l'équation du modèle en précisant ses hypothèses.
- (b) On estime les paramètres du modèle à l'aide du logiciel R en obtenant le listing partiel de la Fig. 2. Peut-on conclure que les paramètres sont différents de zéro ? Pourquoi ?

```
              Estimate Std. Error    t value    Pr(>|t|)
(Intercept) 18.8753139   1.1025428  17.1198019 4.972302e-17
poids       -0.6381624   0.6566083  -0.9719073 3.388683e-01
```

FIGURE 2 – Portion de sortie R avec les coefficients du modèle de régression linéaire de la vitesse  $y$  en fonction du poids  $x$ .

2. On s'intéresse ensuite à la relation entre la puissance  $x'$  (en chevaux-vapeur - ch) et  $y$ , voir Fig. 1, plot de droite. Le listing de la Fig. 3 donne le modèle de régression linéaire.

- (a) Donner les estimations des paramètres du modèle et écrire la droite des moindres carrés.
- (b) Combien de secondes fait gagner chaque cheval-vapeur supplémentaire?
- (c) Donner une mesure de la qualité du modèle. Expliquer.

```
Call:
lm(formula = vitesse ~ puissance)

Residuals:
    Min       1Q   Median       3Q      Max
-2.1766 -0.6975  0.0348  0.6520  4.0972

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  20.556354   0.542424  37.897 < 2e-16 ***
puissance    -0.018458   0.003359  -5.495 5.77e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.282 on 30 degrees of freedom
Multiple R-squared:  0.5016,    Adjusted R-squared:  0.485
F-statistic: 30.19 on 1 and 30 DF,  p-value: 5.766e-06
```

FIGURE 3 – Sortie R avec le modèle de régression linéaire de la vitesse  $y$  en fonction de la puissance  $x'$ .