Université Jean Monnet
Licence 3 Informatique

KERDONCUFF Tanguy
SEBBAN Marc

Correction du TD9 de probabilités

Table des matières

1 TD9 : Tests d'hypothèses

1.1 Exercice I

La tension de sortie d'un transformateur est de 130 volts, d'après ses spécifications. Un ensemble de 40 mesures a donné une moyenne de 128.6 avec un écart-type de 2.1.

1. Tester l'hypothèse que le voltage soit de 130 volts, contre l'hypothèse qu'il soit inférieur, avec un niveau de signification de 5%.

   Solution: On pose $H_0 : \mu = 130$ et $H_a : \mu < 130$.
   On suppose H_0, deux possibilités, soit on réfute H_0 soit on ne peut pas réfuter H_0. Ce n'est pas qu'on accepte H_0 mais bien qu'on ne peut pas réfuter cette hypothèse. On a une analogie avec le principe mathématiques «raisonnement par l'absurde ». On va supposer H_0 par l'absurde et voir si on obtient une absurdité.
   On a $N = 40$ mesures donc $\bar{Y} \hookrightarrow \mathcal{N}(\mu = 130, \sigma = 2.1)$. On va utiliser l'écart type donné expérimentalement car on a pas mieux.
   On va cherche k tel que $P(\bar{Y} \leq k) = 0.05$. Puis comparer cette valeur à la moyenne empirique obtenu.

$$\begin{aligned}
P(\bar{Y} \leq k) = 0.05 &\implies P\left(Z \leq \frac{k-\mu}{\sigma/\sqrt{N}}\right) = 0.05 \\
&\implies P\left(Z \leq \frac{k-130}{2.1/\sqrt{40}}\right) = 0.05 \text{ On lit dans la table que } P(Z \leq -1.65) = 0.05 \\
&\implies \frac{k-130}{2.1/\sqrt{40}} = -1.65 \\
&\implies k = 130 - \frac{2.1}{\sqrt{40}} \times 1.65 \\
&\implies k = 129.45
\end{aligned}$$

Donc si H_0 est vrai alors la probabilité que la réalisation de $\bar{Y}$ soit de moins de $k = 129.45$ est de 0.05. Or on a une réalisation de $\bar{Y}$ de 128.6, ce qui est très peu probable. On a donc une absurdité, on rejet l'hypothèse H_0.

2. Si le voltage descend au-dessous de 128 volts, les circuits situés derrière le transformateur seront détruits. En considérant l'hypothèse nulle $H_0 : \mu = 130$ et l'hypothèse alternative $H_a : \mu = 128$, calculer l'erreur de type II pour la région de rejet déterminée en 1.

   Solution: Erreur de type II : $\beta = P(H_0|H_a)$. Avec $H_a : \mu = 128$.

$$\begin{aligned}
\beta =& P(H_0|H_a) \\
=& P(Z > \frac{129.46 - 128}{2.1/\sqrt{40}} \\
=& P(Z > 4.4) \\
=& 0
\end{aligned}$$

1.2 Exercice II

Un fabriquant d'électroménager produit des réfrigérateurs en 3 couleurs, blanc, jaune et bleu. Sur les 1000 premiers exemplaires vendus, on remarque que 400 sont blancs.

1. Doit-on en conclure que les consommateurs choisissent la couleur au hasard ?

   Solution: Problème intéressant. L'hypothèse nulle est $H_0 : p = \frac{1}{3}$ l'hypothèse opposé est $H_a : p > \frac{1}{3}$. On a $N = 1000$ mesures, on pose $\hat{p} = \frac{1}{N}\sum_{i=1}^{N} C_i$, somme de Bernoulli IID donc Binomiale. Comme on a $N > 30$ on peut supposer $\hat{p} \hookrightarrow \mathcal{N}(\mu = p = \frac{1}{3}, \sigma = \frac{\sqrt{Np(1-p)}}{N} = \sqrt{\frac{2/9}{1000}})$. Ici on prend bien comme moyenne $p = \frac{1}{3}$ car on suppose H_0, de la même manière l'écart type dépend de p et non pas de la réalisation de $\hat{p} = 0.4$.
   Ici on ne nous donne pas la valeur de α qui doit déterminer le seuil de confiance. On fait le raisonnement inverse : au lieu de trouver la valeur de p qui correspond au α on regarder si il est possible que

l'évènement $\hat{p} \geq 0.4$ arrive. Si c'est très improbable, on rejette H_0.

$$\begin{aligned}
P(\hat{p} \geq 0.4) =& P\left(Z \geq \frac{0.4-\mu}{\sigma}\right) \\
=& P\left(Z \geq \frac{0.4-\frac{1}{3}}{\sqrt{\frac{2/9}{1000}}}\right) \\
=& P\left(Z \geq 4.472\right) \\
=& 0
\end{aligned}$$

Pour rappel : $Z_{\alpha=0.01} = 2.33$ et $Z_{\alpha=0.05} = 1.65$. Si on suppose H_0 il est très improbable que l'évènement arrive donc on rejet l'hypothèse H_0.

1.3 Exercice III

Une tour de distillation est censée produire 800 tonnes de produit chimique par jour. Les quantités produites quotidiennement pendant une semaine ont été de 795, 805, 790, 793 et 802 tonnes. On supposera que la production suit une loi normal.

1. Ces données indiquent-elles que la production moyenne quotidienne est inférieure à 800 tonnes, et donc que la tour a un problème (au seuil de 95%) ?

   Solution: On a $H_0 : \mu = 800$ et $H_a : \mu \leq 800$. On suppose H_0. On a les valeurs de l'échantillon, on calcul la moyenne et la variance empirique. $\bar{Y} = \frac{1}{5}(795+805+790+793+802) = 797$. $S^2 = \frac{1}{n-1}\sum_{i=1}^{5}(Y_i-\bar{Y})^2 = 39.5$.
   On va chercher la valeur de k tel que $P(\bar{Y} \leq k) = 0.05$ Soit la valeur limite de la moyenne, en dessous de laquelle on pourra réfuter H_0. Ici cas particulier, on a une variance estimer empiriquement et n ¡ 30. Donc on doit utiliser la loi de Student avec 5-1=4 degré de liberté !

$$\begin{aligned}
P(\bar{Y} \leq 797) =& P\left(T_4 \leq \frac{797-\mu}{\sigma}\right) \\
=& P\left(T_4 \leq \frac{797-800}{\sqrt{\frac{39.5}{5}}}\right) \\
=& P\left(T_4 \leq -1.067\right) \\
=& 0.15(\text{très approximativement}) = p_{valeur}
\end{aligned}$$

   La probabilité que cet évènement arrive n'est pas très rare (15%), vu que le seul est à 5% on ne rejette pas H_0 Dans la table on trouve : $T_{\alpha=0.05} = -2.132$.

1.4 Exercice IV

On souhaite déterminer si un dé est pipé. On le lance 600 fois et on observe les résultats suivants :

Tirage	Frequence
1	89
2	113
3	98
4	104
5	117
6	79

1. Ces données expérimentales suggèrent-elles que le dé est pipé ? Considérer $\alpha = 5\%$.

   Solution: On effectue un test du χ^2 (Khi-deux). $H_0 : \forall i \in [1,6] p_i = \frac{1}{6}$ H_a : il existe$i \in [1,6] p_i \neq \frac{1}{6}$
   On suppose H_0 donc les 6 V.A. ($Y_1, ..., Y_6$ on une moyenne de $n \times p = 100$ On a plus de 2 tests, donc on va utiliser le test du χ^2. On peut l'utiliser, car on a plus de 30 lancer de dés.

$$\chi^2 = \sum_{i=1}^{6} \frac{Y_i - E(Y_i)}{E(Y_i} \hookrightarrow \chi^2_{6-1}$$

$$\chi^2 = \sum_{i=1}^{6} \frac{y_i - E(Y_i)}{E(Y_i)} = \frac{11^2 + 13^2 + 2^2 + 4^2 + 17^2 + 21^2}{100} = 10.4$$

   Plus cette valeur est proche de 0, plus il est possible que les dés soit non pipé. On regarde la valeur limite, celle avec $\alpha = 0.05$ et 5 degré de liberté : 11.07. Donc H_0 n'est pas rejeter. On ne peut pas dire que le dé et pipé.

1.5 Exercice V

On pose à un échantillon de fumeurs et de non-fumeurs la question «Fumer est-il mauvais pour la santé ?». Les résultats obtenus sont les suivants :

	OUI	NON	Sans Opinion	Total
Fumeurs	430	91	16	537
Non-Fumeurs	958	20	20	998
Total	1388	111	36	1535

1. La réponse dépend-elle du fait que l'on soit fumeur ou pas ? Considérer $\alpha = 0.05$

   Solution:
   On prend H_0 : aucune différence entre fumeur et non fumeur et H_a : la réponse dépend du fait que la personne est fumeur.

On suppose H_0, on doit déjà calculer la valeur que devrait avoir le tableau si il n'y avait aucune différence entre les deux groupes. On cherche $E(Y_i)$. Par exemple : $138.8 \times 537/1535 = 485.6$

	OUI	NON	Sans Opinion	Total
Fumeurs	485.6	38.8	12.6	537
Non-Fumeurs	902.4	72.2	23.4	998
Total	1388	111	36	1535

On applique la formule pour chacune des V.A.

$$\chi^2 = \frac{55.6^2}{485.6} + \frac{52.2^2}{38.8} + \frac{3.4^2}{12.6} + \frac{55.6^2}{902.4} + \frac{52.2^2}{72.2} + \frac{3.4^2}{23.4} = 119.2$$

On est très loin des valeurs possibles pour la loi du khi-deux. Il y n'a que 1% de chance que la valeur soit supérieur à 10. On peut donc clairement rejeter l'hypothèse H_0
degré de liberté $ddl = (l-1)(c-1) = 1 \times 2 = 2$

2. Donner une borne à la p-valeur.

Solution: La valeur n'est même pas tabulé... Vraiment trop petit.