Fiche 10 - Statistiques - M1 ingé math
Tests classiques

Emeline Schmisser, `emeline.schmisser@math.univ-lille1.fr`, bureau 314 (bâtiment M3).

Exercice 1 : Test sur la moyenne, variance inconnue
On observe $x_1, \dots, x_n$ qui suivent une loi normale $\mathcal{N}(\mu, \sigma^2)$. Les paramètres μ et σ^2 sont inconnus. On voudrait tester :

$$\begin{cases} H_0 : & \mu = 0 \\ H_1 : & \mu \neq 0 \end{cases}$$

1. Proposer une statistique de test T et donner sa loi sous H_0.
2. Donnez la forme de la région de rejet W.
3. Si $n = 10$, que vaut W pour un niveau $\alpha = 5\%$?
4. Même question si $n = 121$.

Exercice 2 : Boîtes de chocolat
Sur les boîtes de chocolat, on peut lire : contenance 500g. En réalité, la quantité contenue dans les boîtes est une variable gaussienne X. On observe les données suivantes (en grammes) :
490 490 490 492 492 495 497 497 502 505

1. Le fabricant tient-il ses engagements ? On fera le test au niveau $\alpha = 5\%$.

Exercice 3 : Les faiseurs de pluie
La hauteur annuelle des pluies dans la Beauce (en mm) suit une loi normale $\mathcal{N}(600, 100^2)$. Des entrepreneurs, surnommés "faiseurs de pluies" prétendaient pouvoir augmenter de 50 mm le niveau moyen de pluie par l'insémination des nuages au moyen d'iodure d'argent et augmenter ainsi le taux de production. Leur procésé fut mis à l'essai entre 1951 et 1959 et on releva les hauteurs de pluies suivantes :

Année	1951	1952	1953	1954	1955	1956	1957	1958	1959
mm	510	614	780	512	501	534	603	788	650

1. Les entrepreneurs veulent prouver que leur méthode fonctionne (donc que la hauteur de pluie est $> 600mm$). Faire le test, en supposant d'abord que la variance est toujours de 100^2, puis en supposant la variance inconnue. Peuvent-ils convaincre que leur méthode marche ?
2. Les autorités veulent prouver que la méthode ne marche pas (que la hauteur des pluies est $< 650mm$). Faire le test, en supposant d'abord que la variance est égale à 100^2, puis en supposant que la variance est inconnue. Peut-on penser que les faiseurs de pluie sont des charlatans ?

Exercice 4 : Loi normale, bis
On effectue 10 mesures indépendantes sur une population supposée obéir à une loi normale $\mathcal{N}(\mu, \sigma^2)$ et on obtient les résultats suivants :

$$\sum_{i=1}^{n} x_i = 30 \quad \text{et} \quad \sum_{1}^{n} x_i^2 = 117$$

1. Proposer un estimateur de $\hat{\sigma^2}$. Quelle est sa loi ?

2. Construire un test de niveau $\alpha = 5\%$:
$$\begin{cases} H_0 : & \sigma = 1.5 \\ H_1 : & \sigma \neq 1.5 \end{cases}$$
3. Quelle est la probabilité de rejeter H_0 quand l'écart-type vaut 3 ?

Exercice 5 : Rythme cardiaque

On prélève au hasard deux échantillons de personnes, l'un en milieu urbain, l'autre en milieu rural. On évalue, pour chacun des individus, le rythme cardiaque au repos. On obtient les résultats suivants :

	milieu urbain	milieu rural
effectif de l'échantillon	30	24
moyenne du rythme cardiaque	80	77
variance du rythme cardiaque	150	120

On commence par tester l'égalité des variances.

1. Formuler H_0 et H_1.
2. Sous H_0, quelle est la loi asymptotique de
$$T = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} ?$$
3. En déduire la forme de la région critique asymptotique.
4. On fixe le niveau à 5%. On trouve $F_{29,23}(0.975) = 2.25$, $F_{29,23}(0.95) = 1.97$, $F_{23,29}(0.975) = 2.17$ et $F_{23,29}(0.95) = 1.91$. On accepte ou on rejette H_0 ?

On admet désormais que les variances sont égales. On s'intéresse maintenant à l'égalité des moyennes.

5. Formuler H_0 et H_1.
6. Quelle est la loi asymptotique de $n_1\hat{\sigma}_1^2 + n_2\hat{\sigma}_2^2$?
7. Quelle est la loi asymptotique sous H_0 de $\bar{x}_1 - \bar{x}_2$ sous H_0 ?
8. En déduire une statistique permettant de savoir si les moyennes sont égales. Quelle est la loi asymptotique de cette statistique ?
9. Peut-on penser que le fait d'être en milieu rural ou urbain a une influence sur le rythme cardiaque avec un risque de 5% ?

Exercice 6.

Dans une usine, deux machines fabriquent la même pièce. À la fin de la journée, on prélève deux échantillons indépendants de taille respective 100 et 120. Le nombre de pièces non conformes est de 12 pour la première machine et de 16 pour la seconde. On veut savoir si la proportion de pièces défectueuses est la même pour les deux machines ou non.

1. Réaliser un test asymptotique pour un niveau $\alpha = 5\%$.

Exercice 7 : Pasteur contre la génération spontanée

Au 19ème siècle, on pensait que les moisissures apparaissaient spontanément dans l'air. Louis Pasteur a montré qu'au contraire, les germes de moisissures étaient présents dans l'air, et d'autant plus présents que l'air est pollué. Pour cela, il prépare 20 flacons de levure. Il y injecte ensuite de l'air prélevé en ville, à la campagne et au-dessus de la mer de glace. Il observe les résultats suivants :

– Sur les 20 échantillons prélevés en ville, 17 présentent des signes de moisissures.
– Sur les 20 échantillons prélevés à la campagne, 12 présentent des signes de moissisures.
– Sur les 20 échantillons prélevés à la mer de glace (un glacier), 1 seul présente des signes de moisissures.

1. Est-ce que les germes apparaissent par génération spontanée (dans ce cas, la probababilité de développer une moisissure est la même qu'on soit à la montagne, à la campagne ou en ville) ou est-ce qu'ils sont déjà présents dans l'air (et de façon différente en ville, à la campagne et à la montagne) ? On fera le test au niveau $\alpha = 1\%$.

Exercice 8 : Mendel

On cultive 2 variétés de petits pois : des petits pois jaunes et des petits pois ridés. On croise ces deux variétés entre elles. L'hybride obtenu a une apparence normale, ils sont verts et lisses (les gènes qui donnent la couleur jaune et ceux qui donnent l'apparence ridée sont donc tous les deux récessifs). On laisse ensuite l'hybride obtenu s'autofertiliser. Cette fois-ci, on obtient 4 types de petits pois : normaux (verts et lisses), jaunes et lisses, verts et ridés et ridés et jaunes. D'après les lois de Mendel, le nombre de petits pois dans chaque catégorie doit suivre une multinomiale de paramètre (9/16,3/16,3/16,1/16). Les résultats de l'expérience sont :

normaux	jaunes	ridés	jaunes et ridés
670	230	238	62

1. Tester au niveau $\alpha = 10\%$ si les lois de Mendel sont vérifiées.

Exercice 9 : Exercice de révision

Soit X une variable aléatoire de densité $f_\theta(x) = \frac{1}{\theta}x^{\frac{1}{\theta}-1}\mathbf{1}_{[0,1]}(x)$ $(\theta > 0)$. On observe un échantillon $(X_1, \ldots, X_n)$ de variables i.i.d.

1. Préciser le modèle. Est-ce que ce modèle est exponentiel ?
2. Calculer l'estimateur $\hat{\theta}$ du maximum de vraisemblance.
3. Est-ce que cet estimateur est SBVM ?
4. Calculer l'information de Fisher.
   Indication: : on fera un changement de variables et une intégration par parties.
5. Montrer que $\hat{\theta}$ est un estimateur efficace de θ.
6. Quelle est la loi asymptotique de $\hat{\theta}$?
7. Quelle est la loi asymptotique de $1/\hat{\theta}$?

On veut tester $H_0 : \theta = 1$ contre $H_1 : \theta = 1/2$.

8. Donner la forme de la zone de rejet optimale.

On observe seulement (X_1, X_2), et on décide de rejeter H_0 si $X_1X_2 > 3/4$.

9. *Montrer que* $-\ln(X)$ *suit une loi exponentielle de paramètre* $1/\theta$. Quelle est la loi de $-\ln(X_1X_2)$?

La loi Γ de R dépend de deux paramètres : le premier correspond à n et le second à θ. *R nous donne* `pgamma(-log(3/4),2,1)= 0.03423845` *et* `pgamma(-log(3/4),2,1/2)=0.1138577`

10. Quel est le niveau de ce test ?
11. Quelle est la puissance de ce test ?
12. Est-ce que ce test est uniformément plus puissant ?