

Solution de l'exercices 6 de la Série N°2

Exercice 6.

Dans une production il y a une proportion p d'articles défectueux. On prélève 40 articles pour réaliser le test

$$\begin{cases} H_0 : p \leq 0.1 \\ H_1 : p > 0.1 \end{cases},$$

au niveau de signification 0.01. Déterminer le test optimal puis tracer le graphe de sa fonction puissance.

Solution. Ici, la variable aléatoire X prend une valeur 1 (qui correspond à un article défectueux) avec une probabilité p et elle prend la valeur 0 (qui correspond à un article non défectueux). Donc il s'agit d'une v.a qui suit la loi de Bernoulli de paramètre p . La fonction masse de cette v.a discrète X est définie comme suit:

$$\mathbf{P}(X = 1) = p \text{ et } \mathbf{P}(X = 0) = 1 - p.$$

Celle-ci peut être reformuler comme suit:

$$\mathbf{P}(X = x) = p^x (1 - p)^{1-x}, \quad x = 0, 1.$$

Nous allons appliquer la méthode du test de *rapport de vraisemblance monotone*. Soit $p_1 > p_2$ et écrivons

$$\frac{L_{p_1}}{L_{p_2}} = \frac{\prod_{i=1}^{40} p_1^{x_i} (1 - p_1)^{1-x_i}}{\prod_{i=1}^{40} p_2^{x_i} (1 - p_2)^{1-x_i}} = \frac{p_1^t (1 - p_1)^{40-t}}{p_2^t (1 - p_2)^{40-t}}, \text{ avec } t := \sum_{i=1}^{40} x_i.$$

Celle peut être réécrite comme suit:

$$\frac{L_{p_1}}{L_{p_2}} = \left(\frac{1 - p_1}{1 - p_2} \right)^{40} \left(\frac{p_1 (1 - p_2)}{p_2 (1 - p_1)} \right)^t.$$

On pose $b := \left(\frac{1 - p_1}{1 - p_2} \right)^{40} > 0$, et $a := \frac{p_1(1-p_2)}{p_2(1-p_1)} > 1$ (car $p_1 > p_2$), ainsi $t \rightarrow \frac{L_{p_1}}{L_{p_2}}(t) = ba^t$ est une fonction croissante

en t (car $a > 1$). La statistique $T := \sum_{i=1}^{40} X_i$ est une v.a discrete qui suit la loi binomiale de parametre $(40, 0.1)$.

Nous allons enoncer un resultat qui nous permet de l'approcher par une v.a continue.

Proposition 1. Soit $X_1, \dots, X_n$ une suite v.a iid qui suit la loi Bernoulli de parametre p . Supposons que

$$n > 30, \quad np > 5, \quad n(1 - p) > 5.$$

Alors

$$S := \sum_{i=1}^n X_i \rightsquigarrow \mathcal{N}(np, np(1 - p)).$$

Dans notre exercice nous avons $n = 40 > 30$, et $np = 40 \times 0.1 = 4$. Nous allons alors appliquer la proposition 3, pour avoir le test up :

$$\delta = \begin{cases} 1 & \text{si } \sum_{i=1}^{40} x_i \geq c \\ \gamma & \text{si } \sum_{i=1}^{40} x_i = c \\ 0 & \text{si } \sum_{i=1}^{40} x_i < c \end{cases}$$

où c et $0 < \gamma < 1$ sont telles que

$$\mathbf{P}_{p=0.1} \left(\sum_{i=1}^{15} X_i > c \right) + \gamma \mathbf{P}_{p=0.1} \left(\sum_{i=1}^{15} X_i = c \right) = \alpha = 0.01.$$

En d'autres termes

$$1 - \mathbf{P}_{p=0.1} \left(\sum_{i=1}^{15} X_i \leq c \right) + \gamma \mathbf{P}_{p=0.1} \left(\sum_{i=1}^{15} X_i = c \right) = 0.01,$$

ce qui implique que

$$\mathbf{P}_{p=0.1} \left(\sum_{i=1}^{15} X_i \leq c \right) = \gamma \mathbf{P}_{p=0.1} \left(\sum_{i=1}^{15} X_i = c \right) + 0.99 > 0.99.$$

On note que $Y := \sum_{i=1}^{15} X_i \rightsquigarrow \text{Binomial}(15, 0.1)$ (la loi binomiale de paramètre $n = 15$ et $p = 0.1$). De la table statistique, de la loi binomiale, la plus petite c telle que $\mathbf{P}(Y \leq c) > 0.99$ est $c = 5$, pour la quelle $\mathbf{P}(Y \leq 5) = 0.998$. Par conséquent

$$\begin{aligned} \gamma &= \frac{\mathbf{P}(Y \leq 5) - 0.99}{\mathbf{P}(Y = 5)} \\ &= \frac{0.998 - 0.99}{\mathbf{P}(Y \leq 5) - \mathbf{P}(Y \leq 4)} \\ &= \frac{0.998 - 0.99}{0.998 - \mathbf{P}(Y \leq 4)}. \end{aligned}$$

De la table statistique, de la loi binomiale on trouve $\mathbf{P}(Y \leq 4) = 0.987$, donc

$$\gamma = \frac{0.998 - 0.99}{0.998 - 0.987} = 0.72.$$

Ainsi le test optimal upp est

$$\delta = \delta(X_1, \dots, X_{15}) = \begin{cases} 1 & \text{si } \sum_{i=1}^{15} x_i > 5 \\ 0.72 & \text{si } \sum_{i=1}^{15} x_i = 5 \\ 0 & \text{si } \sum_{i=1}^{15} x_i < 5 \end{cases}$$

La fonction puissance du test est

$$\begin{aligned} \pi(p) &= \mathbf{E}_p[\delta], \quad 0 < p < 1 \\ &= 1 \times \mathbf{P}(\delta = 1) + 0.72 \times \mathbf{P}(\delta = 0.72) + 0 \times \mathbf{P}(\delta = 0). \end{aligned}$$

En d'autres termes

$$\pi(p) = \mathbf{P}(T > 5) + 0.72\mathbf{P}(T = 5),$$

où $T := \sum_{i=1}^{15} X_i$ est une v.a binimiale de paramettrer $(15, p)$, avec $0 < p < 1$. Il est claire que

$$\begin{aligned} \pi(p) &= 1 - \mathbf{P}(T \leq 5) + 0.72(\mathbf{P}(T \leq 5) - \mathbf{P}(T \leq 4)) \\ &= 1 - 0.28\mathbf{P}(T \leq 5) - 0.72\mathbf{P}(T \leq 4) \\ &= 1 - 0.28F_p(5) - 0.72F_p(4), \end{aligned}$$

où $F_p(x)$ désigne la fonction de répartition de la loi binomiale de paramètrer $(15, p)$. Explicitement la fonction puissance est

$$\pi(p) = 1 - 0.28 \sum_{k=0}^5 C_{15}^k p^k (1-p)^{15-k} - 0.72 \sum_{k=0}^4 C_{15}^k p^k (1-p)^{15-k},$$

pour $0 < p < 1$. Le graphe de fonction puissance est donné par la figure Fig.1.

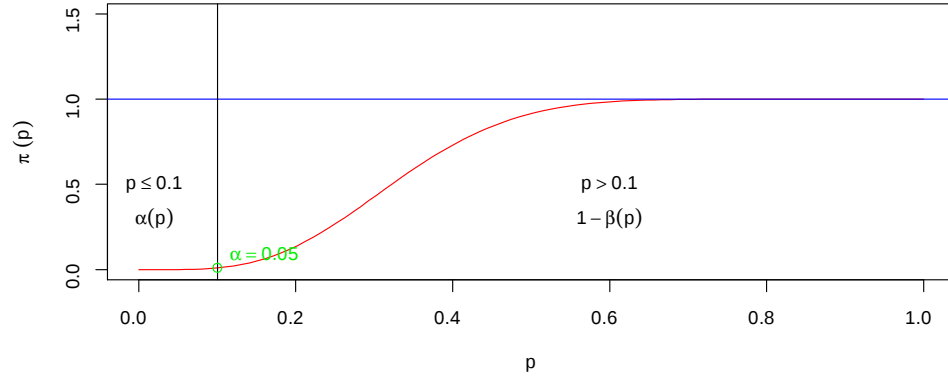


Fig.1

Le code R de cette figure est:

```
-----
f<-function(x){1-0.28*pbinom(5,15,x)-0.72*pbinom(4,15,x)}
x<-seq(0,1,length=100)
plot(x,f(x),type="l",col="red",ylim=c(0,1.5),xlab=expression(p),ylab=expression(pi~(p)))
abline(h=1,col="blue")
abline(v=0.1)
points(0.1,0.01,col="green2")
text(0.16,0.1,expression(alpha==0.05),col="green2")
text(0.6,0.3,expression(1-beta(p)))
text(0.02,0.3,expression(alpha(p)))
text(0.02,0.5,expression(p<=0.1))
text(0.6,0.5,expression(p>0.1))
-----
```