



3.4.8

Théorie:

H_0 hypothèse nulle

H_1 hypothèse alternative

Hypothèses: $\begin{cases} H_0 \text{ et } H_1 \text{ incompatibles} \\ \text{l'une des deux est vraie} \end{cases}$

H_0 : "vrai par défaut"

$\rightarrow$ pour rejeter H_0 il faut une bonne raison

R_1 : zone de rejet

T statistique de test, la loi de T n'est pas la même selon que H_0 est vraie ou fausse

Niveau: α tq $P(R_1 | H_0) \leq \alpha$

Test de niveau 5% = tq le probas de rejeter H_0 à tort soit 5%

$H_0 = "p = \frac{1}{2}"$ pile qu'on lance 100 fois

$H_1 = "p \neq \frac{1}{2}"$

$T = \text{nb de pile}$

$p = \text{probas de pile}$

On rejette si $T \leq 2$ ou $T \geq 98$

$P(\text{rejeter} | H_0)$ très très faible

Puissance: $\inf P(\text{rejeter} | p)$ $p \neq \frac{1}{2}$

$\inf \{ P(\text{rejet}) , \theta \in H_1 \}$

θ paramètre $H_1 = \{ \theta \text{ tq } H_1 \text{ vraie} \}$

erreur de type 1: rejeter H_0 à tort $\leadsto$ on veut α petit

erreur de type 2: accepter H_0 à tort $\leadsto$ on veut grande puissance

Exercice 3

1) $X_1, \dots, X_N$ indep, de m loi: $\mathcal{B}(p)$

Alors, $T = X_1 + \dots + X_N \leadsto \mathcal{B}(N, p)$

2) paramètre inconnu: p

H_0 : " $p = p_0$ "

H_1 : " $p \neq p_0$ "

Statistique de test

Pour construire le teste, il suffit de choisir une zone de rejet R_n

On veut: Niveau $\alpha \leq 5\%$

i.e $P(R_n | H_0) \leq 5\%$

i.e $P(R_n | p = p_0) \leq 5\%$

Forme de R_n : on va dire qu'on rejette H_0 si T est

Si H_0 vraie $p = p_0$
 $T \sim \mathcal{B}(n, p_0)$
 T prend une valeur proche de $N_{p_0} = E[T]$
ce serait très surprenant que T soit très éloigné de N_{p_0}

R_n : " $|T - N_{p_0}| \geq c$ " avec c à choisir.

On veut $P(|T - N_{p_0}| \geq c | p = p_0) \leq 5\% \iff P(T \leq N_{p_0} - c) + P(T \geq N_{p_0} + c) \leq 5\%$

Si $N_{p_0} = 90 \iff P(T_0 \leq 90 - c) + 1 - P(T_0 \leq 90 + c - 1) \leq 5\%$

Pour simplifier, on demande $P(T_0 \leq 90 - c) \leq 2,5\% \longrightarrow c = 7$ convient
et $1 - P(T_0 \leq 90 + c) \leq 2,5\%$

$\Rightarrow P(T_0 \leq 90 + c) \geq 97,5\% \rightarrow c = 6$ convient

$P(T_0 \leq 83) \approx 2,06\%$

$P(T_0 \leq 95) \approx 97,63\%$

Bilan: On prend $c=7$

Rejet si $|T-90| \geq 7$

i.e si $T \leq 83$ ou $T \geq 97$

96 fonctionne aussi

3] Vu notre test, si $T=84$ on ne rejette pas H_0 :
On ne peut pas conclure que $p=p_0$
On considère donc que $p=p_0$

Autre réponse à question 2:

On rejette si $T \leq 84$

$$P(\text{rejet} | H_0) = P(T_0 \leq 84) \approx 3.99\%$$

Test de niveau $\alpha = 5\%$

Si on a répondu cela à 2]:

2] Si $T=84$ on rejette H_0 on conclut que $p \neq p_0$

Exercice 4

On lance 50 fois le dé

H_0 : "le dé est équilibré"

H_1 : "le dé est truqué de Sarah"

Statistique de test: N

Sous H_0 on s'attend à ce que la valeur de N
soit proche $\frac{50}{3} \approx 17$

Sous H_1 on s'attend à ce que N soit + grand

Zone de rejet de la forme: " $N \geq c$ " avec $c \geq \frac{50}{3}$

ce qui serait surprenant si H_0 est vraie
beaucoup + raisonnable si H_1 vraie.

On veut niveau 1% i.e $P(\text{rejet} | H_0) \leq 1\%$
i.e $P(N \geq c | H_0) \leq 1\%$
i.e $P(N < c | H_0) > 99\%$
i.e $P(N \leq c-1 | H_0) \geq 99\%$
i.e $P_{\text{acc}}(N \leq c-1) \geq 0,99$
i.e $c-1 \geq 25$
i.e $c \geq 26$ On choisit $c=26$
On rejette si $N \geq 26$

2) Puissance: $\inf_{\theta \in (H_1)} P(\text{rejet})$
= proba de rejeter si θ est vrai
= $P_{\text{acc}}(N \geq 26)$
= $1 - P_{\text{acc}}(N \leq 25)$
= $1 - 0,556 = 0,444$

Si le dé est truqué, on a 44,4% de chance de s'en apercevoir.

If the cube is truced, there's 44% chance to notice it.

3) Si $N = 26$, Jade rejette H_0 : elle accuse Sarah d'avoir menti.

Exercice 8

1) X nombre de pannes sur une journée: $X \sim P(\lambda_0)$ Donc $E(X) = \lambda_0$
 $E(X) \leq 2$. Donc on choisit $\lambda_0 = 2$

2) Pour $i \in [1, 30]$ notons X_i le nb de pannes le jour i

Notons $Z = X_1 + \dots + X_{30}$. Supposons $X_1, \dots, X_{30}$ indép.
Alors $Z \sim P(60)$

3) 4) $P(|Z - E[Z]| > c) \leq \frac{\text{Var}(Z)}{c^2}$
 $P(Z < E[Z] - c) + P(Z > E[Z] + c) \leq \frac{\text{Var}(Z)}{c^2}$
 $P(Z \notin [E[Z] - c, E[Z] + c]) \leq \frac{\text{Var}(Z)}{c^2}$
 $P(Z \in [\underbrace{E[Z] - c}_a, \underbrace{E[Z] + c}_b]) \geq 1 - \frac{\text{Var}(Z)}{c^2}$

On va choisir c tq $1 - \frac{V(Z)}{c^2} \geq 0,95$

$$\Leftrightarrow \frac{V(Z)}{c^2} \leq 0,05 = \frac{1}{20}$$

$$\Leftrightarrow c^2 \geq 20 V(Z)$$

$$\Leftrightarrow c \geq \sqrt{20} \sigma(Z) = \sqrt{20} \sqrt{60} \\ = 20\sqrt{3} \approx 35$$

5] H_0 : "la fréquence des pannes est inchangée"
 H_1 : "elle a changé"
Statistique de test: S

Sous H_0 : $S \sim P(60)$

On rejette si $S \notin [a, b]$ de question 4
i.e si $S \in [25, 95]$
i.e si $S < 25$ ou $S > 95$

C'est un test de niveau 5% par 4]

6] $S = 30$ on ne rejette pas

Exercice 7

$X_1, \dots, X_n$, indép, même loi esp μ , variance V

a] B.T à $X_1 + \dots + X_n$ d'espérance $n\mu$
de variance nV (car indép)

$$P(|X_1 + \dots + X_n - n\mu| \geq n\varepsilon) \leq \frac{nV}{(n\varepsilon)^2} = \frac{V}{n\varepsilon^2}$$

$$b) P(X_1 + \dots + X_n \notin]n\mu - n\varepsilon, n\mu + n\varepsilon[) \leq \frac{V}{n\varepsilon^2} \\ P(X_1 + \dots + X_n \in \underbrace{]n\mu - n\varepsilon, n\mu + n\varepsilon[}_{A(n, \varepsilon)}) \geq 1 - \frac{V}{n\varepsilon^2}$$

Bien pour un test où on connaît μ (sous H_0)

c) On repart de a) qu'on écrit

$$P(\mu \in]X_1 + \dots + X_n - n\varepsilon, X_1 + \dots + X_n + n\varepsilon[) \leq \frac{\nu}{n\varepsilon^2}$$

$$P(\mu \in]\underbrace{\frac{X_1 + \dots + X_n}{n} - \varepsilon, \frac{X_1 + \dots + X_n}{n} + \varepsilon}[) \geq 1 - \frac{\nu}{n\varepsilon^2}$$

$\hat{T}$ intervalle de confiance pour μ
(niveau risque) $\frac{\nu}{n\varepsilon^2}$