**ITI -** Stat - Examen Final 2022/2023 - **3 h**

**Sauf avis contraire, on considèrera dans les exercices un risque de première espèce $\alpha = 5\%$.**

## Exercise 1 — CPCD — 3 points

1. Donnez, à partir d'un échantillon i.i.d $(X_1, \ldots, X_n)$ dont la loi parente est une loi binomiale de paramètre inconnu $p$, l'intervalle de confiance de Wilson de ce paramètre $p$.
2. Dans quel cas recommanderiez vous d'utiliser cette approximation.

## Exercise 2 — La mine — 3 points

Pour étudier la teneur en fer de 400 kg de minerai, on a prélevé au hasard un échantillon de 20 prises de 1,6 kg sur lesquelles on a mesuré la teneur en fer. On note $X_i$ la variable aléatoire de la teneur en fer de l'échantillon $i \in \{1, \ldots, 20\}$ suivant la loi normale $\mathcal{N}(\mu, \sigma^2)$, $\bar{X}$ la moyenne empirique de ces 20 échantillons et $\hat{S}_{n-1}$ l'écart-type empirique (sans biais). On a observé $\bar{X} = 0,2486$ et $\hat{S}_{n-1} = 0,0028$.

1. Un expert a un a priori sur $\sigma^2$. Il suppose que $\sigma^2 = 0,000009$, vérifiez en construisant un intervalle de confiance à $0,9$ sur $\sigma^2$ que nos données sont compatible avec cet a priori sur la variance.
2. Construisez un intervalle de confiance à $0,95$ sur $\mu$ en supposant $\sigma = 0,003$.
3. On suppose qu'on ne connaît pas $\sigma^2$. Construisez l'intervalle de confiance à $0,9$ sur $\mu$.

## Exercise 3 — Les douaniers — 4 points

Pour étudier la teneur en fer de 400 kg de minerai, on prélève au hasard un échantillon de 16 prises de 1,6 kg sur lesquelles nous avons mesuré la teneur en fer. Normalement, le minerait provient de la mine officielle du gouvernement et la teneur en fer des échantillons suit une loi normale $\mathcal{N}(\mu_1, \sigma^2)$, avec $\mu_1 = 0,5$. Mais il existe aussi du minerai de contrebande pour lequel la teneur en fer des échantillons suit une loi normale $\mathcal{N}(\mu_2, \sigma^2)$, avec $\mu_2 = 0,45$. Nous connaissons la valeur de l'écart-type $\sigma = 0,08$, qui est le même pour les deux mines.

1. Nous avons un échantillon de 16 observations d'une de ces deux mines avec comme teneur en fer moyenne : $\bar{X} = 0,465$. A votre avis, de quelle mine proviennent ces échantillons ?
2. Nous avons un autre échantillon de taille 16 d'une même mine avec comme teneur en fer moyenne : $\bar{X} = 0,51$. Nous savons que cet échantillon provient soit de la mine officielle du gouvernement, soit d'une autre mine avec une teneur moyenne en fer plus forte, toujours avec la même variance. A votre avis, de quelle mine proviennent ces échantillons ?

## Exercise 4 — Raie lait (poisson liquide ?) — 5 points

Soit $\theta$ un réel strictement positif. On dira qu'une variable aléatoire X suit une loi de Rayleigh de paramètre $\theta > 0$ si et seulement si sa densité est $f_\theta$, qui est définie par :

$$f_\theta(x) = \frac{x}{\theta} exp\left(-\frac{x^2}{2\theta}\right) \quad \text{si } x \geqslant 0, \text{ et 0 sinon}$$

On notera alors $X \sim R(\theta)$. On considère $n$ variables aléatoires $X_1, X_2, ..., X_n$ indépendantes et de même loi de Rayleigh de paramètre $\theta > 0$, supposé inconnu.

1. Déterminer l'estimateur du maximum de vraisemblance de $\theta$, qu'on le notera $\hat{\theta}$.
2. Cet estimateur est-il efficace ? Est-il sans biais ? Déterminer son espérance et sa variance.
3. Démontrer que
$$\frac{\sqrt{n}(\hat{\theta} - \theta)}{\theta} \quad \underset{n \to \infty}{\overset{(\mathcal{L})}{\longrightarrow}} \quad \mathcal{N}(0, 1)$$
4. Supposant que $n$ est grand, construire un intervalle de confiance bilatéral pour le paramètre $\theta$ au niveau de confiance $(1 - \alpha)$ avec $\alpha \in ]0, 1[$.

## Exercise 5 — Tenue de combat — 2 points

Montrer que le loi uniforme sur $[\theta, \theta+1]$ de densité $f_\theta$ avec

$$f_\theta(x) = \begin{cases} 1 & \text{si } x \in [\theta, \theta+1] \\ 0 & \text{sinon} \end{cases}$$

est à rapport de vraisemblance monotone (on pourra considérer deux cas (en supposant $\theta_1 \leq \theta_2$) : $\theta_2 \leq \theta_1 + 1$ et $\theta_1 + 1 < \theta_2$ ).

## Exercise 6 — Fish — 3 points

On suppose que le nombre $X$ d'absences par semaine dans une entreprise suit une loi de Poisson de paramètre $\lambda$ avec

$$p(k) = P(X = k) = \frac{\lambda^k}{k!} \exp^{-\lambda}$$

On dispose des données suivantes, où $n_x$ est le nombre de semaines pour lesquelles on a relevé $x$ absences :

| $x$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $n_x$ | 9 | 6 | 5 | 4 | 1 |

On suppose que ces données peuvent être considérées comme une réalisation d'un échantillon i.i.d. $X_1, \ldots, X_n$ de taille n = 25, de variable parente $X$. On notera $T$ la statistique $\sum_{i=1}^{n} X_i$. On cherche à tester les deux hypothèses simples :

$$H_0 : \lambda = \lambda_0$$
$$H_1 : \lambda = \lambda_1 \ (> \lambda_0).$$

1. Montrer que la loi de poisson est une distribution à rapport de vraisemblance monotone.
2. En déduire que la région critique optimale s'exprime en fonction de la statistique $T$.
3. En déduire la région critique lorsque $\lambda_0 = 1$ et le résultat du test avec les données de l'exercice.

**Tables de la loi normale :** Cette table nous donne les valeurs de $t$ telle que $P(X \leq t)$ lorsque $X$ suit une loi normale centrée réduite

$$\Pi(t) = P(X \leq t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}\, dx \quad \text{et} \quad \Pi(-t) = 1 - \Pi(t).$$

| t | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1.0 | 0.8413 | 0.8438 | 0.8461 | 0.8485 | 0.8508 | 0.8531 | 0.8554 | 0.8577 | 0.8599 | 0.8621 |
| 1.1 | 0.8643 | 0.8665 | 0.8686 | 0.8708 | 0.8729 | 0.8749 | 0.8770 | 0.8790 | 0.8810 | 0.8830 |
| 1.2 | 0.8849 | 0.8869 | 0.8888 | 0.8907 | 0.8925 | 0.8944 | 0.8962 | 0.8980 | 0.8997 | 0.9015 |
| 1.3 | 0.9032 | 0.9049 | 0.9066 | 0.9082 | 0.9099 | 0.9115 | 0.9131 | 0.9147 | 0.9162 | 0.9177 |
| 1.4 | 0.9192 | 0.9207 | 0.9222 | 0.9236 | 0.9251 | 0.9265 | 0.9279 | 0.9292 | 0.9306 | 0.9319 |
| 1.5 | 0.9332 | 0.9345 | 0.9357 | 0.9370 | 0.9382 | 0.9394 | 0.9406 | 0.9418 | 0.9429 | 0.9441 |
| 1.6 | 0.9452 | 0.9463 | 0.9474 | 0.9484 | 0.9495 | 0.9505 | 0.9515 | 0.9525 | 0.9535 | 0.9545 |
| 1.7 | 0.9554 | 0.9564 | 0.9573 | 0.9582 | 0.9591 | 0.9599 | 0.9608 | 0.9616 | 0.9625 | 0.9633 |
| 1.8 | 0.9641 | 0.9649 | 0.9656 | 0.9664 | 0.9671 | 0.9678 | 0.9686 | 0.9693 | 0.9699 | 0.9706 |
| 1.9 | 0.9713 | 0.9719 | 0.9726 | 0.9732 | 0.9738 | 0.9744 | 0.9750 | 0.9756 | 0.9761 | 0.9767 |

**Tables de la loi de Student :** Cette table nous donne les valeurs de $t$ telle que $\mathbb{P}(T > t)$ lorsque $T$ suit une loi de student à $\nu$ degrés de liberté

| $\nu$ | 0,10 | 0,05 | 0,025 | 0,01 | 0,005 | 0,001 |
|---|---|---|---|---|---|---|
| 1 | 3,078 | 6,314 | 12,706 | 31,821 | 63,657 | 318,313 |
| 2 | 1,886 | 2,920 | 4,303 | 6,965 | 9,925 | 22,327 |
| 3 | 1,638 | 2,353 | 3,182 | 4,541 | 5,841 | 10,215 |
| 4 | 1,533 | 2,132 | 2,776 | 3,747 | 4,604 | 7,173 |
| 5 | 1,476 | 2,015 | 2,571 | 3,365 | 4,032 | 5,893 |
| 6 | 1,440 | 1,943 | 2,447 | 3,143 | 3,707 | 5,208 |
| 7 | 1,415 | 1,895 | 2,365 | 2,998 | 3,499 | 4,782 |
| 8 | 1,397 | 1,860 | 2,306 | 2,896 | 3,355 | 4,499 |
| 9 | 1,383 | 1,833 | 2,262 | 2,821 | 3,250 | 4,296 |
| 10 | 1,372 | 1,812 | 2,228 | 2,764 | 3,169 | 4,143 |
| 11 | 1,363 | 1,796 | 2,201 | 2,718 | 3,106 | 4,024 |
| 12 | 1,356 | 1,782 | 2,179 | 2,681 | 3,055 | 3,929 |
| 13 | 1,350 | 1,771 | 2,160 | 2,650 | 3,012 | 3,852 |
| 14 | 1,345 | 1,761 | 2,145 | 2,624 | 2,977 | 3,787 |
| 15 | 1,341 | 1,753 | 2,131 | 2,602 | 2,947 | 3,733 |
| 16 | 1,337 | 1,746 | 2,120 | 2,583 | 2,921 | 3,686 |
| 17 | 1,333 | 1,740 | 2,110 | 2,567 | 2,898 | 3,646 |
| 18 | 1,330 | 1,734 | 2,101 | 2,552 | 2,878 | 3,610 |
| 19 | 1,328 | 1,729 | 2,093 | 2,539 | 2,861 | 3,579 |
| 20 | 1,325 | 1,725 | 2,086 | 2,528 | 2,845 | 3,552 |
| 21 | 1,323 | 1,721 | 2,080 | 2,518 | 2,831 | 3,527 |
| 22 | 1,321 | 1,717 | 2,074 | 2,508 | 2,819 | 3,505 |
| 23 | 1,319 | 1,714 | 2,069 | 2,500 | 2,807 | 3,485 |
| 24 | 1,318 | 1,711 | 2,064 | 2,492 | 2,797 | 3,467 |
| 25 | 1,316 | 1,708 | 2,060 | 2,485 | 2,787 | 3,450 |
| 26 | 1,315 | 1,706 | 2,056 | 2,479 | 2,779 | 3,435 |
| 27 | 1,314 | 1,703 | 2,052 | 2,473 | 2,771 | 3,421 |
| 28 | 1,313 | 1,701 | 2,048 | 2,467 | 2,763 | 3,408 |
| 29 | 1,311 | 1,699 | 2,045 | 2,462 | 2,756 | 3,396 |
| 30 | 1,310 | 1,697 | 2,042 | 2,457 | 2,750 | 3,385 |

**Tables de la loi du chi 2 :** Cette table nous donne les valeurs de $t$ telle que $\mathbb{P}(X > t)$ lorsque $X$ suit une loi du chi 2 à $\nu$ degrés de liberté

| $\nu$ | 0.3 | 0.2 | 0.1 | 0.05 | 0.025 | 0.01 | 0.005 | 0.001 |
|---|---|---|---|---|---|---|---|---|
| 1 | 1.0742 | 1.6424 | 2.7055 | 3.8415 | 5.0239 | 6.6349 | 7.8794 | 10.8276 |
| 2 | 2.4079 | 3.2189 | 4.6052 | 5.9915 | 7.3778 | 9.2103 | 10.5966 | 13.8155 |
| 3 | 3.6649 | 4.6416 | 6.2514 | 7.8147 | 9.3484 | 11.3449 | 12.8382 | 16.2662 |
| 4 | 4.8784 | 5.9886 | 7.7794 | 9.4877 | 11.1433 | 13.2767 | 14.8603 | 18.4668 |
| 5 | 6.0644 | 7.2893 | 9.2364 | 11.0705 | 12.8325 | 15.0863 | 16.7496 | 20.5150 |
| 6 | 7.2311 | 8.5581 | 10.6446 | 12.5916 | 14.4494 | 16.8119 | 18.5476 | 22.4577 |
| 7 | 8.3834 | 9.8032 | 12.0170 | 14.0671 | 16.0128 | 18.4753 | 20.2777 | 24.3219 |
| 8 | 9.5245 | 11.0301 | 13.3616 | 15.5073 | 17.5345 | 20.0902 | 21.9550 | 26.1245 |
| 9 | 10.6564 | 12.2421 | 14.6837 | 16.9190 | 19.0228 | 21.6660 | 23.5894 | 27.8772 |
| 10 | 11.7807 | 13.4420 | 15.9872 | 18.3070 | 20.4832 | 23.2093 | 25.1882 | 29.5883 |
| 11 | 12.8987 | 14.6314 | 17.2750 | 19.6751 | 21.9200 | 24.7250 | 26.7568 | 31.2641 |
| 12 | 14.0111 | 15.8120 | 18.5493 | 21.0261 | 23.3367 | 26.2170 | 28.2995 | 32.9095 |
| 13 | 15.1187 | 16.9848 | 19.8119 | 22.3620 | 24.7356 | 27.6882 | 29.8195 | 34.5282 |
| 14 | 16.2221 | 18.1508 | 21.0641 | 23.6848 | 26.1189 | 29.1412 | 31.3193 | 36.1233 |
| 15 | 17.3217 | 19.3107 | 22.3071 | 24.9958 | 27.4884 | 30.5779 | 32.8013 | 37.6973 |
| 16 | 18.4179 | 20.4651 | 23.5418 | 26.2962 | 28.8454 | 31.9999 | 34.2672 | 39.2524 |
| 17 | 19.5110 | 21.6146 | 24.7690 | 27.5871 | 30.1910 | 33.4087 | 35.7185 | 40.7902 |
| 18 | 20.6014 | 22.7595 | 25.9894 | 28.8693 | 31.5264 | 34.8053 | 37.1565 | 42.3124 |
| 19 | 21.6891 | 23.9004 | 27.2036 | 30.1435 | 32.8523 | 36.1909 | 38.5823 | 43.8202 |
| 20 | 22.7745 | 25.0375 | 28.4120 | 31.4104 | 34.1696 | 37.5662 | 39.9968 | 45.3147 |
| 21 | 23.8578 | 26.1711 | 29.6151 | 32.6706 | 35.4789 | 38.9322 | 41.4011 | 46.7970 |
| 22 | 24.9390 | 27.3015 | 30.8133 | 33.9244 | 36.7807 | 40.2894 | 42.7957 | 48.2679 |
| 23 | 26.0184 | 28.4288 | 32.0069 | 35.1725 | 38.0756 | 41.6384 | 44.1813 | 49.7282 |
| 24 | 27.0960 | 29.5533 | 33.1962 | 36.4150 | 39.3641 | 42.9798 | 45.5585 | 51.1786 |
| 25 | 28.1719 | 30.6752 | 34.3816 | 37.6525 | 40.6465 | 44.3141 | 46.9279 | 52.6197 |
| 26 | 29.2463 | 31.7946 | 35.5632 | 38.8851 | 41.9232 | 45.6417 | 48.2899 | 54.0520 |
| 27 | 30.3193 | 32.9117 | 36.7412 | 40.1133 | 43.1945 | 46.9629 | 49.6449 | 55.4760 |
| 28 | 31.3909 | 34.0266 | 37.9159 | 41.3371 | 44.4608 | 48.2782 | 50.9934 | 56.8923 |
| 29 | 32.4612 | 35.1394 | 39.0875 | 42.5570 | 45.7223 | 49.5879 | 52.3356 | 58.3012 |
| 30 | 33.5302 | 36.2502 | 40.2560 | 43.7730 | 46.9792 | 50.8922 | 53.6720 | 59.7031 |
| 31 | 34.5981 | 37.3591 | 41.4217 | 44.9853 | 48.2319 | 52.1914 | 55.0027 | 61.0983 |
| 32 | 35.6649 | 38.4663 | 42.5847 | 46.1943 | 49.4804 | 53.4858 | 56.3281 | 62.4872 |
| 33 | 36.7307 | 39.5718 | 43.7452 | 47.3999 | 50.7251 | 54.7755 | 57.6484 | 63.8701 |
| 34 | 37.7954 | 40.6756 | 44.9032 | 48.6024 | 51.9660 | 56.0609 | 58.9639 | 65.2472 |
| 35 | 38.8591 | 41.7780 | 46.0588 | 49.8018 | 53.2033 | 57.3421 | 60.2748 | 66.6188 |
| 36 | 39.9220 | 42.8788 | 47.2122 | 50.9985 | 54.4373 | 58.6192 | 61.5812 | 67.9852 |
| 37 | 40.9839 | 43.9782 | 48.3634 | 52.1923 | 55.6680 | 59.8925 | 62.8833 | 69.3465 |
| 38 | 42.0451 | 45.0763 | 49.5126 | 53.3835 | 56.8955 | 61.1621 | 64.1814 | 70.7029 |
| 39 | 43.1053 | 46.1730 | 50.6598 | 54.5722 | 58.1201 | 62.4281 | 65.4756 | 72.0547 |
| 40 | 44.1649 | 47.2685 | 51.8051 | 55.7585 | 59.3417 | 63.6907 | 66.7660 | 73.4020 |

# Formulaire

— fréquences $\widehat{f_i} = \frac{n_i}{n}$ où $n$ est le nombre total d'observations et $n_i$ le nombre d'observant de la modalité $i$

— moyenne : $\overline{x} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \hat{f}_i x_i$

— variance empirique : $\widehat{\sigma}_x^2 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(x_i - \overline{x})^2$

— espérance d'une variable aléatoire discrète : $\mathbb{E}(X) = \displaystyle\sum_i x_i \mathbb{P}(x_i)$.

— espérance d'une variable aléatoire continue de densité $f(x)$ : $\mathbb{E}(X) = \displaystyle\int x f(x) dx$.

— variance d'une variable aléatoire $X$ : $Var(X) = \mathbb{E}\Big(\big(X - \mathbb{E}(X)\big)^2\Big)$.

— quantile (ou fractiles) à l'ordre $p$, $\forall p \in [0,1], \widehat{\Phi}_p$ telle que $\widehat{\mathbb{P}}(X \leq \widehat{\Phi}_p) = p$

— les quartiles :

  — $\widehat{\Phi}_{\frac{1}{4}} = \widehat{Q}_1$, telle que $\widehat{F}(\widehat{Q}_1) = \frac{1}{4}$,

  — $\widehat{\Phi}_{\frac{1}{2}} = \widehat{Q}_2 = \widehat{M}$, telle que $\widehat{F}(\widehat{M}) = \frac{1}{2}$,

  — $\widehat{\Phi}_{\frac{3}{4}} = \widehat{Q}_3$, telle que $\widehat{F}(\widehat{Q}_2) = \frac{3}{4}$.

— covariance : $c_{xy} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})$ et corrélation : $\mathrm{cor}(x,y) = \dfrac{c_{xy}}{\sqrt{\widehat{\sigma}_x^2 \widehat{\sigma}_y^2}}$

— probabilité conditionnelle : $\mathbb{P}(X = x_i | Y = y_j) = \dfrac{\mathbb{P}(X = x_i, Y = y_j)}{\mathbb{P}(Y = y_j)}$

— espérance conditionnelle : $\mathbb{E}[Y|X = a] = \displaystyle\sum_{i=1}^{n} y_i \mathbb{P}(Y = y_i | X = a)$

— La loi des grands nombres $\bar{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i \quad \underset{n\to\infty}{\longrightarrow} \quad \mathbb{E}(X)$

Estimateurs :

— Le risque : $R_{\widehat{\theta}}(\theta) = \mathbb{E}\big((\widehat{\theta}(X_1, \ldots, X_n) - \theta)^2\big)$

  — Le biais d'un estimateur : $\mathbb{E}(\widehat{\theta}) - \theta$

  — La variance d'un estimateur : $\mathbb{E}\Big(\big(\widehat{\theta} - \mathbb{E}(\widehat{\theta})\big)^2\Big)$

— Soit $X$ une variable aléatoire de distribution $f_\theta(x)$ et $(X_1, \ldots, X_n)$ une famille de $n$ variables aléatoires i.i.d. ayant pour loi parente la loi de $X$.

  — Vraisemblance :

$$L(\theta, X_1, \ldots, X_n) = \prod_{i=1}^{n} f_\theta(X_i)$$

  Log vraisemblance :

$$\ell(\theta, X_1, \ldots, X_n) = log\, L(\theta, X_1, \ldots, X_n) = \sum_{i=1}^{n} log\, f_\theta(X_i)$$

— L'information de Fisher : $I_n(\theta) = Var\Big(\dfrac{\partial \ell(\theta, X_1, \ldots, X_n)}{\partial \theta}\Big) = -\mathbb{E}\left(\dfrac{\partial^2 \ell(\theta, X_1, \ldots, X_n)}{\partial \theta^2}\right)$

— Pour $\theta \in \mathbb{R}^p$, L'information de Fisher est la matrice $p \times p$

$$I_n(\theta) = \mathbb{E}\big(\nabla \ell(\theta, X_1, \ldots, X_n) \nabla \ell(\theta, X_1, \ldots, X_n)^\top\big)$$

— Borne de Cramer Rao

  — pour un estimateur sans biais de $\theta$

$$BCR(\theta) = \frac{1}{I_n(\theta)} \;\leq\; \mathrm{Var}\big(\widehat{\theta}(X_1, \ldots, X_n)\big)$$

  — Pour un estimateur sans biais d'une fonction de $\theta$ : $\mathbb{E}(\widehat{u}(\theta)) = u(\theta)$

$$BCR(\theta) = \frac{u'(\theta)}{I_n(\theta)} \;\leq\; \mathrm{Var}\big(\widehat{u}(\theta(X_1, \ldots, X_n)\big)$$

— pour un estimateur biaisé de biais $B$

$$BCR(\theta) = \frac{1 + B'(\theta)}{I_n(\theta)} \ \leq \ \mathrm{Var}\big(\widehat{\theta}(X_1, \ldots, X_n)\big)$$

— L'estimateur max de vraisemblance $\widehat{\theta}_{MV}$ d'un paramètre $\theta$ est :
  — Asymptotiquement sans biais

$$\widehat{\theta}_{MV} \underset{n\to\infty}{\longrightarrow} \theta^\star$$

  — Asymptotiquement efficace

$$Var(\widehat{\theta}_{MV}) \underset{n\to\infty}{\longrightarrow} I_n^{-1}(\theta^\star)$$

  — Asymptotiquement normal

$$\sqrt{n}(\widehat{\theta}_{MV} - \theta^\star) \underset{n\to\infty}{\longrightarrow} \mathcal{N}\big(0, I_1^{-1}(\theta^\star)\big)$$

  note : $I_n(\theta^\star) = nI_1(\theta^\star)$

Variables aléatoires et lois
— Soit $Y \sim \mathcal{N}(0,1)$ une variable aléatoire normale centrée réduite.
— Soit $Y_1, Y_2, ..., Y_n$ un échantillon de $n$ réalisation i.i.d. de cette variable aléatoire.
— La loi du $\chi^2$ : On appelle loi du $\chi^2$ à $n$ degrés de libertés la loi de la variable aléatoire $Z_n = \sum_{i=1}^n Y_i^2$
— La loi de student : On appelle loi de student à $n$ degrés de libertés la loi de la variable aléatoire $T_n$

$$T_n = \frac{N}{\sqrt{\frac{X_n}{n}}} \qquad \begin{matrix} N \sim \mathcal{N}(0,1) \\ X_n \sim \chi^2_n \end{matrix}$$

Tests
— Régions de décision :
  Région d'acceptation : l'ensemble $\overline{W}$ des réalisations $(x_1, \ldots, x_n)$ pour lesquelles on garde $H_0$
  Région critique : l'ensemble $W$ des réalisations $(x_1, \ldots, x_n)$ pour lesquelles on rejette $H_0$
— Mesures de performances des tests
  — Risque de première espèce ou niveau de signification du test $\alpha$ : probabilité de rejeter $H_0$ alors que $H_0$ est vraie

$$\alpha = \mathbb{P}(D = 1 | H_0) = \int_{(x_1,\ldots,x_n)\in W} L(x_1, \ldots, x_n, H_0)\, dx_1, \ldots, dx_n$$

  — Risque de seconde espèce $\beta$ : probabilité de garder $H_0$ alors que $H_1$ est vraie

$$\beta = \mathbb{P}(D = 0 | H_1) = \int_{(x_1,\ldots,x_n)\in \overline{W}} L(x_1, \ldots, x_n, H_1)\, dx_1, \ldots, dx_n$$

  — puissance d'un test : $1 - \beta$ la probabilité de rejeter hypothèses nulle avec raison
  — $p$-valeur :
— Tests du rapport de vraisemblance

$$W = \left\{ x_1, \ldots, x_n \mid \frac{L(x_1, \ldots, x_n, H_1)}{L(x_1, \ldots, x_n, H_0)} > k \right\}$$

— Le principe de Neyman–Pearson : pour un $\alpha$ fixé, trouver la fonction de décision qui minimise $\beta$ (ou qui maximise $1 - \beta$ la puissance du test)
— Le théorème de Neyman–Pearson : pour un test paramétrique de deux hypothèses simples, le test du rapport de vraisemblance

$$W = \left\{ x_1, \ldots, x_n \mid \frac{L(x_1, \ldots, x_n, H_1)}{L(x_1, \ldots, x_n, H_0)} > k \right\}$$

tel que $k$ soit fixé de sorte que

$$\alpha = \int_{(x_1,\ldots,x_n)\in W} L(x_1, \ldots, x_n, H_0)\, dx_1, \ldots, dx_n$$

est optimal au sens du principe de Neyman–Pearson (parmi tous les tests de risque $\alpha$, c'est celui de puissance maximale).