Statistiek (WISB263)

Final exam

June 28, 2023

Schrijf uw naam op elk in te leveren vel. Schrijf ook uw studentnummer op blad 1.
`(The exam is a CLOSED-book exam: students can bring only two A4-sheets with personal notes. The use of the statistical tables is allowed. The scientific calculator is also allowed).`

The maximum number of points is 100.
Points distribution: 20–24–26–30.

1. (a) [8pt] Let Y_1 and Y_2 be two i.i.d. random variables such that $Y_i \sim \text{Poi}(\lambda)$ for $i \in \{1,2\}$ and $\lambda \in \mathbb{R}_+$, i.e.,

$$\mathbb{P}(Y_i = k) = e^{-\lambda}\frac{\lambda^k}{k!} \text{ with } k \in \mathbb{N}_0.$$

Show that $Y_1 + Y_2 \sim \text{Poi}(2\lambda)$.

(b) [8pt] For the random variable $Y \sim \text{Poi}(100)$ find an approximated value of the probability $\mathbb{P}(Y > 120)$.

(c) [4pt] Show that:

$$\lim_{n\to\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \frac{1}{2}$$

2. (a) [8pt] Assume that the outcome of an experiment is the realization of a single random variable Y, whose distribution depends on an unknown parameter $\theta \in \mathbb{R}$. A 80% confidence interval for θ has the form $[X-1, X+2]$. Determine a decision rule and the rejection region for testing:

$$\begin{cases} H_0: & \theta = 5, \\ H_1: & \theta \neq 5. \end{cases}$$

at $\alpha = 0.2$ level of significance.

(b) Let Y_1 and Y_2 be two independent random variables, such that $Y_i \sim \text{Uniform}[0, \theta]$, for $i \in \{1,2\}$. We want to test:

$$\begin{cases} H_0: & \theta = 1, \\ H_1: & \theta > 1. \end{cases}$$

and we reject H_0 when $\max(Y_1, Y_2) > c$.

(i) [8pt] Find c so that the test has significance level $\alpha = \frac{19}{100}$.

(ii) [8pt] Which is the power function of the test (as a function of the value θ_1 of H_1)?

3. Consider a multinomial distribution with probability mass function:

$$\mathbb{P}(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3) = \frac{m!}{y_1! y_2! y_3!} p_1^{y_1} p_2^{y_2} p_3^{y_3},$$

with $\sum_{i=1}^{3} p_i = 1$ and $\sum_{i=1}^{3} y_i = m$.

(a) [8pt] Show that the maximum likelihood estimate $\hat{p}_i$ of p_i is $\frac{y_i}{m}$ with $i \in \{1,2,3\}$.

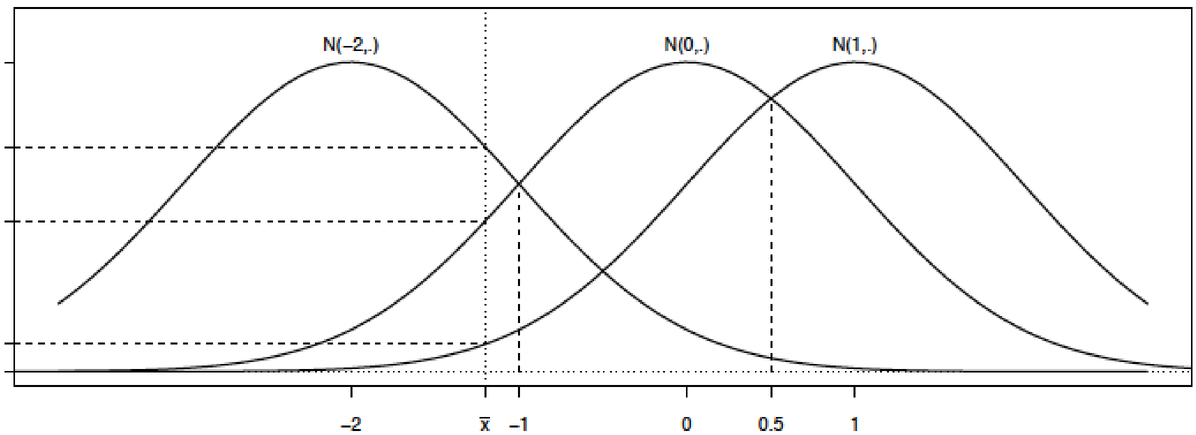
(b) [8pt] It is suspected that $p_1 = p_2 = p$, where $0 < p < 1$. Show that the maximum likelihood estimate $\hat{p}$ of p is then $\frac{y_1+y_2}{2m}$.

(c) [10pt] Find the generalized likelihood ratio test statistic for comparing the two models of point (a) and point (b). State its asymptotic distribution and find the rejection region for a test at $\alpha = 0.05$ level of significance.

4. Consider the sample $\mathbf{X} = \{X_1, \dots X_n\}$ of i.i.d. random variables such that $X_i \sim N(\theta, \sigma^2)$ with σ^2 **known** and $\theta \in \Theta$, where the parameter space is the discrete set $\Theta = \{-2, 0, 1\}$.

   (a) [4pt] Show that $\bar{X} := 1/n \sum_{i=1}^{n} X_i$ is a sufficient statistic for θ and that the likelihood $L(\theta; \mathbf{X})$ can be factorized in $L(\theta; \mathbf{X}) = h(\mathbf{X}) g_\theta(\bar{X})$, with:

$$h(\mathbf{X}) = (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X})^2\right\}, \qquad g_\theta(\bar{X}) = \exp\left\{-\frac{n}{2\sigma^2}(\bar{X} - \theta)^2\right\}$$



In the figure above the function $g_\theta(y)$ is plotted for the three possible values of the parameter θ.

   (b) [8pt] Find the maximum likelihood estimator $\hat{\theta}_{MLE}$ of θ.

   (c) [8pt] Find the probability mass function of $\hat{\theta}_{MLE}$.

   (d) [4pt] Is $\hat{\theta}_{MLE}$ a biased estimator?

   (e) [6pt] Find the most powerful test for testing:

$$\begin{cases} H_0: & \theta = 0, \\ H_1: & \theta = 1, \end{cases}$$

   at α level of significance. Can you say something about the rejection region of this test?