Computational Statistics

DM
DEPARTAMENTO DE MATEMÁTICA
TÉCNICO LISBOA

BSc in Applied Mathematics and Computation
MSc in Applied Mathematics and Computation
MSc in Data Science and Engineering

1st semester – 2025/2026
18/12/2025 – 19:00

2nd Test Duration: 45 minutes

Justify all the answers properly!

Single group 10 values

1. Let X be the proportion of students preferring online classes in a survey. Assume that the random variable X has cumulative distribution function (cdf) $F(x) \equiv P(X \leq x) = \int_0^x 12\, t^2 (1-t)\, dt,\ \ 0 < x < 1$.

(a) Estimate $S(x) = P(X > x)$ by proposing a simple Monte Carlo method via sampling the Uniform $(0,1)$ distribution, with random sample denoted by $u_1, \ldots, u_m$. (2.0)

(b) For nine values of x, compare the estimates obtained based on (a) for $m = 1000$ with those computed numerically, which are below stored in the vectors $\widehat{S(x)}$ and $S(x)$, respectively. (2.0)

x	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\widehat{S(x)}$	0.997	0.978	0.931	0.852	0.742	0.607	0.461	0.322	0.214
$S(x)$	0.996	0.973	0.916	0.821	0.688	0.525	0.348	0.181	0.052

2. Let Y denote the number of insurance claims (counts) where the claim size X (in euros) matters. Let $\mathscr{D} = \{(y_1, x_1), \ldots, (y_n, x_n)\}$ be a realization of a random sample for n days of the Poisson-Normal model characterized by $Y|X \sim \text{Poisson}(\eta\delta^X)$, $\eta > 0$, $\delta > 0$, and $X \sim \text{Normal}(\mu, \tau^{-1})$, $\mu \in I\!R$, $\tau = \frac{1}{\sigma^2} > 0$. For $\theta = (\eta, \delta, \mu, \tau)$, the non-informative prior distribution given by $h(\eta, \delta, \mu, \tau) \propto (\eta\delta\tau)^{-1}$.

(a) Show that the joint posterior distribution of θ, given $\mathscr{D}$, is proportional to (2.0)

$$\eta^{\sum_{i=1}^n y_i - 1} \delta^{\sum_{i=1}^n x_i y_i - 1} \tau^{\frac{n}{2}-1} \exp\Big(-\eta \sum_{i=1}^n \delta^{x_i} - \frac{\tau}{2} a(x,\mu)\Big).$$

where $a(x,\mu) = \sum_{i=1}^n (x_i - \bar{x})^2 + n(\mu - \bar{x})^2$.

(b) Using the Gibbs sampler for the joint posterior distribution of θ, presented in (a), find the full conditional posterior distributions of θ, identifying those that are well-known distributions. (2.0)

(c) Introducing the parameterization $\eta = e^{\beta_0}$ and $\delta = e^{\beta_1}$ in the joint posterior distribution of θ, presented in (a), find the full conditional distribution of β_1, denoted by $h(\beta_1|\mathscr{D}, \beta_0, \mu, \tau)$, and comment if the adaptive rejection algorithm can be used to generate that distribution in the Gibbs sampler. (2.0)