Statistical Computing: Homework 2

Due on March 24 (Thursday) 8:30

1. Consider a Weibull model with the pdf

$$f(x;\theta,\beta) = \frac{\beta}{\theta}\left(\frac{x}{\theta}\right)^{\beta-1} e^{-(x/\theta)^{\beta}}, \quad \theta > 0, \beta > 0, x > 0.$$

   Let $\boldsymbol{\eta} = (\theta, \beta)'$. Use Monte Carlo method to evaluate the following quantities for the cases with parameter values $\boldsymbol{\eta} = (1, 0.5)$ and $\boldsymbol{\eta} = (1, 2)$.

   (a) kurtosis: $E\left(\frac{X-EX}{\sqrt{var(X)}}\right)^4$.

   (b) Fisher information matrix (2×2):

$$I(\boldsymbol{\eta}) = -E\left[\frac{\partial^2}{\partial\boldsymbol{\eta}\partial\boldsymbol{\eta}'}\log f(X;\boldsymbol{\eta})\right].$$

   (c) Use numerical integration by deterministic method (such as "cuhre" in R) to get the answer in (b).

   (d) Use R comment "rweibull" to generate data of size $n = 100$. Compute the empirical Fisher information based on your data, and compare it to your answer in (b).

   Use MH mathods to draw weibull samples in (a)(b) for obtaining the Monte Carlo approximation. Your write-up should include the following:

   - sampling scheme
   - Monte Carlo estimate and the corresponding Monte Carlo estimation s.e.

2. Use Monte Carlo method to investigate the power function of the JB test (for testing normality) against the following non-Gaussian models:

   - Alternative Model I: skew-normal distribution (hw1 Problem #4) with skew parameter θ:

$$f(x;\theta) = 2\,\phi(x)\,\Phi(\theta x), \quad x \in R, \quad \theta > 0.$$

   - Alternative Model II: t(ν) with df $\nu = 1/\theta$, $0 < \theta \leq 1$.

When $\theta = 0$, both models reduce to the standard normal distribution (H_0). A larger θ indicates more deviations from normality.

Sketch the power function $\pi(\theta)$ for one alternative model (at your own choice) at a fine grid $\mathcal{G}$ on $\theta \in [0, 1]$. In particular $\pi(0)$ corresponds to the type I error for testing normality (H_0). Since the testing power also depends on the sample size n, explore the power function for two cases: (a) data with a small sample size $n = 20$, (b) data with a large sample size $n = 100$. Provide the Monte Carlo error for approximating $\pi(\theta)$ (use large enough m to ensure that the MC s.e. of $\hat{\pi}(\theta_i)$ is less than 0.01).

Your simulation may take the following flow:

- Set the sample size n, the number of MC replicates m, and the type I error $\alpha = 0.05$.
- For each $\theta_i \in \mathcal{G}$,

  (1) do $j = 1 : m$,

  - (a) generate iid random variable from the alternative model with the given θ_i;
  - (b) compute the test statistic and obtain the testing result (reject/ not-reject);

  end do

  (2) compute the rejection rate among m Monte Carlo replicates as $\hat{\pi}(\theta_i)$
- Plot $\hat{\pi}(\theta_i)$ v.s. θ_i to sketch the power function.