SOLUTIONS MAS472: 2016-17

1. (i) (Bookwork)

$$\hat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{X_i \leq x}$$

(ii) (Bookwork)

$$n\hat{F}_n(x) \sim Bin(n, F(x))$$

and so $\hat{F}_n(x) \sim \frac{1}{n}Bin(n, F(x))$

(iii) (Routine)

$$\gamma(\hat{F}_n) = \frac{\mathbb{E}_{\hat{F}_n}(X - \mathbb{E}_{\hat{F}_n}(X))^3}{\left(\mathbb{E}_{\hat{F}_n}(X - \mathbb{E}_{\hat{F}_n}(X))^2\right)^{3/2}}$$

$$= \frac{\frac{1}{n}\sum(X_n - \bar{X}_n)^3}{\left(\frac{1}{n}\sum(X_n - \bar{X}_n)^2\right)^{3/2}}$$

(iv) (Bookwork) By sampling with replacement from $\{X_1, \ldots, X_n\}$

(v) (Unseen) For $b = 1, \ldots, B$

- Sample $X_1*, \ldots, X_n^* \sim \hat{F}_n(\cdot)$
- Calculate $T_b = \gamma(\mathbf{X}^*)$

Then estimate the standard error using

$$\text{std. err}(\hat{\gamma}) = \left(\frac{1}{B-1}\sum_{b=1}^{B}(T_b - \bar{T})^2\right)^{1/2}$$

where $\bar{T} = \frac{1}{B}\sum_{b=1}^{B} T_b$

A 95% CI could be formed either by looking at the 2.5th and 97.5th percentiles of the $\{T_b\}_{b=1}^{B}$ or by

$$\hat{\gamma} \pm 1.96\,\text{std. err}(\hat{\gamma})$$

2. (i) (a) (Unseen) $H_0 : \beta = 0$

(b) (Unseen) We've used a randomisation test. Under H_0, there is no relationship between x and y, and so it should not matter if we permute the elements of x.

(c) (Routine) The test gives an estimated p-value of $31/1000 = 0.031$, and so we reject H_0 at the 5% level, but not at the 1% level.

(d) (Routine) The plot suggests that the distributional assumptions made about ϵ for the t-test may not hold, in particular, the constant variance assumption looks questionable. The randomization test, because it does not rely on a constant variance assumption, may be preferred to the classical t-test.

(e) (Unseen) If there were only 5 observations, then there are $5! = 120$ different permutations. If the observed ordering gives the largest values of $\hat{\beta}$, then the reverse ordering would give $-\hat{\beta}$, and so the smallest possible p-value would be $2/120$.

(ii) (Unseen) For $n = 1, \ldots, N$

- Simulate $\tilde{y}_i \sim N(0, x_i^2)$ for $i = 1, \ldots, 100$
- Calculate $\hat{\beta}_n = \arg\min_\beta \sum_{i=1}^{100} \left(\frac{\tilde{y}_i - \beta x_i}{x_i}\right)^2$

Find

$$p = \frac{1}{N}\sum_{n=1}^{N} \mathbb{I}_{|\hat{\beta}_m| \geq |\hat{\beta}_{obs}|}$$

and reject H_0 if $p < 0.01$, otherwise conclude there is insufficient evidence against H_0.

(iii) (a) (Bookwork) If $L(\alpha, \beta|\mathbf{x}) = \prod_{i=1}^{n} f(x_i; \alpha, \beta)$ is the likelihood function, then the profile likelihood for α is

$$L_p(\alpha) = \max_\beta L(\alpha, \beta|\mathbf{x})$$

(b) (Bookwork) If $l_p(\alpha) = \log L_p(\alpha)$ is the profile log-likelihood, then we define the profile deviance to be

$$D_p(\alpha) = 2\{l_p(\hat{\alpha}) - l_p(\alpha)\}$$

For n large, at the true value of α, $D_p(\alpha) \sim \chi_1^2$, and so a 95% confidence interval for α takes the form

$$C_\alpha = \{\alpha : D_p(\alpha) \leq c_{0.05}\}$$

where $c_{0.05}$ is the 95th percentage point of the χ_1^2 distribution.

3. (i) (a) (Bookwork) Anything sensible about computing using pseudo-random numbers (a deterministic sequence) that have similar properties to the random numbers. The most common approach is to use a congruential generator of some kind.

(b) (Routine)

$$F(x) = \int_0^x f(y)dy = \int_0^y 3x^3 e^{-3x^3} dx$$
$$= 1 - e^{-x^3}$$

Set $U = F(X) = 1 - e^{-X^3}$ so that

$$X = (-\log(1-U))^{1/3}$$

is a random draw from f when $U \sim U[0,1]$.

(c) (Unseen) An unbiased estimator of $\mathbb{E}X$ is

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i = \frac{1}{n}\sum_{i=1}^n g(U_i)$$

Then

$$\mathbb{V}\text{ar}(\bar{X}_n) = \frac{1}{n}\mathbb{V}\text{ar}(g(U_i)) = \frac{0.1054}{n}$$

We know by the central limit theorem that $\bar{X}_n$ will be approximately normally distributed for large n, and so a 95% confidence interval for $\mathbb{E}X$ is

$$\bar{X}_n \pm 1.96\sqrt{\frac{0.105}{n}}$$

which has width $2 \times 1.96\sqrt{\frac{0.105}{n}}$.

Thus if we set

$$n = \frac{0.105}{\left(\frac{10^{-3}}{2\times 1.96}\right)^2} = 1619619$$

we will get a CI with the required width.

(d) (Unseen) We can use the antithetic variables estimator with $2n$ calls to g

$$\tilde{X}_{2n} = \frac{1}{2}\left(\frac{1}{n}\sum_{i=1}^n g(U_i) + \frac{1}{n}\sum_{i=1}^n g(1-U_i)\right)$$

Then

$$\mathbb{V}\text{ar}(\tilde{X}_{2n}) = \frac{1}{2}\left(\mathbb{V}\text{ar}(\bar{X}_n) + \mathbb{C}\text{ov}(\bar{X}_n, \bar{X}'_n)\right)$$

where $\bar{X}'_n = \frac{1}{n}\sum_{i=1}^n g(1-U_i)$. We can see that

$$\mathbb{C}\text{ov}(\bar{X}_n, \bar{X}'_n) = \frac{\mathbb{C}\text{ov}(g(U), g(1-U))}{n}$$

and so we get

$$\begin{aligned}\mathbb{V}\text{ar}(\tilde{X}_{2n}) &= \frac{1}{2n}\left(\mathbb{V}\text{ar}(g(U)) + \mathbb{C}\text{ov}(g(U), g(1-U))\right) \checkmark \\ &= \frac{1}{2n}(0.1054 - 0.1050) = \frac{0.0004}{2n} \checkmark\end{aligned}$$

Setting $2 \times 1.96\sqrt{\frac{0.0004}{2n}} = 10^{-3}$ suggests we now only need a sample size of at least $2n = 6147$ samples. ✓✓

4. (i) (a) (Bookwork) Let $w(x) = \frac{f(x)}{g(x)}$. To estimate S, for $i = 1, \ldots, N$

- Simulate $X_i \sim g(\cdot)$
- Set $w_i = w(X_i)$✓

Then estimate S by

$$\hat{S} = \frac{1}{n} \sum w_i X_i^2 ✓$$

(b) (Routine) First find the mode.

$$\frac{d}{dx} h(x) = \frac{3}{x} - 4x^3 ✓$$

which has a minimum at

$$m = \left(\frac{3}{4}\right)^{1/4} ✓$$

A second order Taylor expansion of $h(x)$ about m is

$$h(x) = h(m) + \frac{1}{2} h''(m)(x-m)^2 ✓$$

where

$$h''(x) = -\frac{3}{x^2} - 12x^2$$

and so

$$h''(m) = -12 \left(\frac{4}{3}\right)^{1/2} = -13.9 ✓$$

and so

$$h(x) = h(m) - \frac{1}{2}(x-m)^2/0.0721$$

So the optimal importance sampling distribution is

$$N\left(\left(\frac{3}{4}\right)^{1/4}, \frac{1}{12}\left(\frac{3}{4}\right)^{1/2}\right) = N(0.93, 0.072) ✓✓$$

(ii) (a) (Unseen) The probability a $N(0, \sigma^2)$ rv is greater than a is $1 - \Phi(a/\sigma) = \Phi(-a/\sigma)$✓. Thus, the amount of time we have to wait until we see an acceptance has a Geometric($\Phi(-a/\sigma)$)✓ distribution, which has mean

$$1/\Phi(-a/\sigma) ✓$$

(b) (Routine)

$$M = \sup_x \frac{f(x)}{g(x)} = \sup_x \frac{c\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{x^2}{2\sigma^2}}\mathbb{I}_{x>a}}{\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}}$$

$$= \sup_x c\exp\left(-\frac{1}{2\sigma^2}(2x\mu - \mu^2)\right)\mathbb{I}_{x>a}$$

This is maximised where $2x\mu - \mu^2$ is minimized, which is at $x = a$ (given that $\mu > 0$ and $x > a$).

So

$$M = c\exp\left(-\frac{1}{2\sigma^2}(2a\mu - \mu^2)\right)$$

So the rejection algorithm is

- Simulate $X \sim N(\mu, \sigma^2)$
- Accept X with probability $\frac{f(X)}{Mg(X)}$

(c) (Unseen) The efficiency of the rejection algorithm is determined by the acceptance rate, which is $1/M$. So we should choose μ to minimize M (and thus maximize $1/M$).

M is minimized when we maximize $2a\mu - \mu^2$, which occurs at $\mu = a$