Week 5

STA 380

2. Consider the integral:

$$\theta = \int_0^{10} x \ln(x) dx$$

(a) Use the simple Monte Carlo estimator to **estimate** the expected value, where the importance function is the $Uniform(0, 10)$ distribution. Call this $\hat{\theta}$.

(b) Write an exact expression for $\mathbb{V}ar[\hat{\theta}]$. **You do not need to evaluate this integral by hand.** Then, compute this integral using `integrate()` in `R`.
Disclaimer: you need to know how to integrate for the tests; we also just want you to learn the `integrate()` function as well. You will not be using the `integrate()` function in `R` for the term tests or exam.

(c) Recall in lecture, we said the Monte Carlo estimate for the variance of the estimator for the integral is:

$$\widehat{\mathbb{V}ar[\hat{\theta}]} = \frac{\hat{\sigma}^2}{n} = \frac{\sum_{i=1}^n \left[g(x_i) - \overline{g(x)}\right]^2}{n^2}$$

Using this equation, compute $\widehat{\mathbb{V}ar[\hat{\theta}]}$ in `R`. Use $n = 10^4$.

(d) Compare the MC estimator against the theoretical variance of $\hat{\theta}$. Use `all.equal()`, which will report the mean relative difference.

(a)

$$\theta = 10 \int_0^{10} \frac{1}{10} \cdot x \cdot \ln(x)\, dx$$

$$= 10\, E_{U(0,10)} [x \cdot \ln(x)]$$

samples from $U \overset{iid}{\sim} \text{Uniform}(0, 10)$.

by SLLN, the sample MC estimator is

$$\hat{\theta} = 10 \times \frac{1}{n} \sum_{i=1}^{n} u_i \cdot \ln(u_i)$$

$Var(X+Y) = Var(X) + Var(Y)$

$Var(Y) = E(Y^2) - (E(Y))^2$

$E(Y) = \int y \cdot f(y)\, dy$

(b).

$$Var[\hat{\theta}] = Var\left[\frac{10}{n} \sum_{i=1}^{n} u_i \cdot \ln(u_i)\right]$$

$$= \frac{100}{n^2} Var\left[\sum_{i=1}^{n} u_i \cdot \ln(u_i)\right]$$

$$= \frac{100}{n^2} \sum_{i=1}^{n} Var[u_i \cdot \ln(u_i)]$$

$$= \frac{100}{n^2} \sum_{i=1}^{n} \left(E\left[(u_i \cdot \ln(u_i))^2\right] - \left(E[u_i \cdot \ln(u_i)]\right)^2 \right)$$

$$= \frac{100}{n^2} \sum_{i=1}^{n} \left(\int_0^{10} (u_i \cdot \ln(u_i))^2 \cdot \frac{1}{10} \, du_i - \left(\int_0^{10} u_i \cdot \ln(u_i) \cdot \frac{1}{10} \, du_i \right)^2 \right)$$

$$C$$

$$= \frac{100}{n^2} \, n \cdot C = \frac{100}{n} \left(\int_0^{10} (u_i \cdot \ln(u_i))^2 \cdot \frac{1}{10} du_i - \left(\int_0^{10} u_i \cdot \ln(u_i) \cdot \frac{1}{10} du_i \right)^2 \right)$$

(C)

$$\widehat{\mathbb{V}ar[\hat{\theta}]} = \frac{\hat{\sigma}^2}{n} = \frac{\sum_{i=1}^{n} \left(\left[g(x_i) - \overline{g(x)} \right] \right)^2}{n^2}$$

$n = 10^4$

Uniform (0,1)

u ← runif(n)

```
n <- 10^4

u <- runif(n, 0, 10)

mc_est <- sum((10* log(u)*u - mean(10* log(u)* u))^2)/n^2
```

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} 10 \, u_i \cdot \ln(u_i)$$

$g(x)$

(b) integrate( )

$fun1 = \int_0^{10} (\mu_i \ln(\mu_i))^2 \cdot \frac{1}{10}$ $E(Y^2)$

$fun2 = \int_0^{10} (\mu_i \ln(\mu_i)) \cdot \frac{1}{10}.$ $E(Y)$

```
fun1 <- function(x){log(x)^2*x^2/10}
fun2 <- function(x){log(x)*x/10}

ex2 <- integrate(fun1, 0, 10)
ex_sq <- integrate(fun2, 0, 10)
exact <- 100 * (ex2$value - (ex_sq$value)^2) / n
```

$$\frac{100}{n}\left(\int_0^{10} (\mu_i \cdot \ln(\mu_i))^2 \cdot \frac{1}{10} d\mu_i - \left(\int_0^{10} \mu_i \cdot \ln(\mu_i) \cdot \frac{1}{10} d\mu_i\right)^2\right)$$

Test 1

Q3.

$z_1, z_2 \cdots z_n$ i.i.d. $z_i \sim N(0,1)$ for all $i = 1, 2 \cdots n$.

chi-squared distribution $\sum_{i=1}^{n} z_i^2 \sim \chi_n^2$

Use rnorm() ... $n = 10^4$ samples from the $\chi^2_{(30)}$.

R codes

```
n <- 10^4

values <- numeric(n)            # vector.

for (i in 1:n){                 # loop

values[i] <- sum(rnorm(30, mean=0, sd=1)^2)

}
```

SumRow

$$\begin{array}{c} 1 \\ 2 \\ 3 \\ \vdots \\ 10^4 \end{array}\begin{bmatrix} Z_{11} & Z_{12} & \cdots & Z_{130} \\ & & & \\ & & & \\ & & & \\ & & & \end{bmatrix}$$

sum ^{2}.

Q4. $f_i(x)$ density $i=1,2..k$.

$f_1(x)$ ...... $f_k(x)$.

π_1, ... π_k weights.

$$f(x) := \sum_{i=1}^{k} \pi_i f_i(x)$$

where $0 \le \pi_i \le 1$ & $\sum_{i=1}^{k} \pi_i = 1$

(a) $f(x)$ non-negative

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

① Non-negative.

Since $f_i(x)$ is valid density $\Rightarrow$ $0 \le f_i(x)$

And $0 \le \pi_i \le 1$

$\Rightarrow 0 \le \sum_{i=1}^{n} \pi_i f_i(x)$

$\int(A+B)$
$= \int A + \int B$

② $\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{\infty} \sum_{i=1}^{n} \pi_i f_i(x)\,dx$

$= \sum_{i=1}^{n} \int_{-\infty}^{\infty} \pi_i \cdot f_i(x)\,dx.$

$= \sum_{i=1}^{n} \pi_i \underbrace{\int_{-\infty}^{\infty} f_i(x)\,dx}_{1} = \sum_{i=1}^{n} \pi_i = 1$

(b) suppose X_i has f_i.

$$\mu_i = E[X_i] = \int_{-\infty}^{\infty} x \cdot f_i(x)\, dx.$$

$$\sigma_i^2 := V[X_i]$$

WTS $\Rightarrow$ $\sigma^2 : V[X] = \sum_{i=1}^{k} \pi_i \sigma_i^2 + \sum_{i=1}^{k} \pi_i (\mu_i - \mu)^2$

$$V[X] = E[(X-\mu)^2]$$

$$= \int_{-\infty}^{\infty} (x-\mu)^2 \cdot \sum_{i=1}^{k} \pi_i \cdot f_i(x)\, dx$$

$$= \int_{-\infty}^{\infty} (x^2 - 2x\mu + \mu^2) \cdot \sum_{i=1}^{k} \pi_i \cdot f_i(x)\, dx$$

$$= \sum_{i=1}^{k} \pi_i \int_{-\infty}^{\infty} (x^2 - 2x\mu + \mu^2)\, f_i(x)\, dx$$

$-2\mu \int_{-\infty}^{\infty} x f_i(x)\, dx$ E

$$= \sum_{i=1}^{k} \pi_i \left[\int_{-\infty}^{\infty} x^2 \cdot f_i(x)\, dx - \int_{-\infty}^{\infty} 2x\mu \cdot f_i(x)\, dx + \int_{-\infty}^{\infty} \mu^2 f_i(x)\, dx \right]$$

$$= \sum_{i=1}^{k} \pi_i \left[E[X^2 | i] - 2\mu\mu_i + \mu^2 \right]$$

$$= \sum_{i=1}^{k} \pi_i \left[V[X|i] + (E[X|i])^2 - 2\mu\mu_i + \mu^2 \right]$$

$$= \sum_{i=1}^{k} \pi_i \left[\sigma_i^2 + \mu_i^2 - 2\mu\mu_i + \mu^2 \right]$$

$$= \sum_{i=1}^{k} \pi_i (\sigma_i^2 + (\mu_i - \mu)^2)$$

$$= \sum_{i=1}^{k} \pi_i \sigma_i^2 + \sum_{i=1}^{k} \pi_i (\mu_i - \mu)$$