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)

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

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

samples from $U \sim$ Uniform $(0, 10)$. $\quad$ i.i.d.

by SLLN, the sample MC estimator is

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

(b)

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

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

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

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

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

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

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

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

$$\underbrace{\phantom{\int_0^{10} (u_i \cdot \ln(u_i))^2 \cdot \frac{1}{10} \, du_i - \left(\int_0^{10} u_i \cdot \ln(u_i) \cdot \frac{1}{10} \, du_i \right)^2}}_{C}$$

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

(C)

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

$$n = 10^4$$

Uniform $[0,1]$

$u \leftarrow runif(n)$

$n \leftarrow 10\verb|^|4$

$u \leftarrow runif(n, 0, 10)$

$mc\_est \leftarrow sum\left(\left(10 * \log(u) * u - mean(10 * \log(u) * u) \right)\verb|^|2 \right) / n\verb|^|2$

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

$$g(x)$$

(b) integrate ( )

$$\text{fun1} = \int_0^{10} \left(u_i \, \ln(u_i) \right)^2 \cdot \frac{1}{10} \quad \leftarrow \quad E(Y^2)$$

$$\text{fun2} = \int_0^{10} \left(u_i \, \ln(u_i) \right) \cdot \frac{1}{10} \quad \leftarrow \quad E(Y)$$

```
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
```

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

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

$$\sum_{i=1}^{n} z_i^2 \sim X_n^2$$

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

R codes

$n \leftarrow 10^4$ 🖊

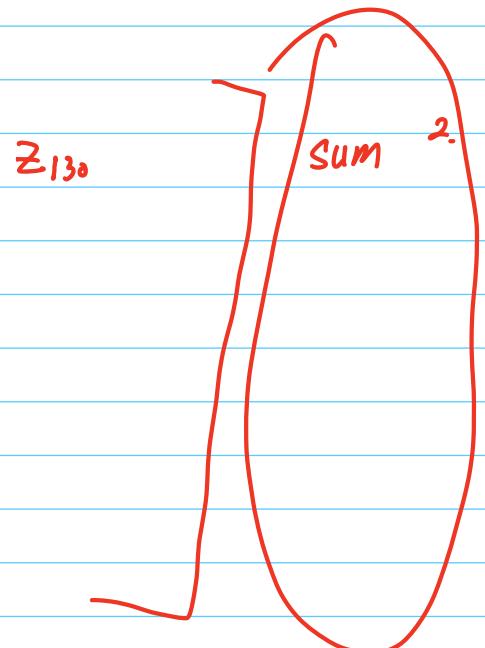# vector.   values $\leftarrow$ numeric (n)

loop   for (i in 1:n) {

values [i] $\leftarrow$ sum (rnorm (30, mean = 0, sd = 1)^2)

}

Sum Row

1	z_{11}	$z_{12} \cdots$		z_{130}	Sum^2
2					
3					
⋮					
10^4					

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

$f_1(x) \ldots\ldots f_k(x).$

$\pi_1, \quad \cdots\ldots \pi_k \quad$ weights.

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

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

(a) $f(y)$ 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 $\quad 0 \le \pi_i \le 1$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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