

Statistics 406 Lab 2
GSI: Adam Rothman
Tuesday September 19

1. Write an R program to exemplify the following property of the variance

For any random variable X , with $\text{var}(X) < \infty$

$$\text{var}(aX) = a^2\text{var}(X) \quad (1)$$

using a random variable $X \sim \text{Normal}(0,1)$, with $a = 2$, with 10000 repetitions.

```
a <- 2
reps <- 10000

## Generate a single N(0,1) random draw, 10000 times
Z <- rnorm(reps)

## Find the variance of these draws
var(Z)

## Find the variance of a times each draw
var(a*Z)

# Verify that the var(a*Z) = a^2 * var(Z)
a^2*var(Z)
```

Here is a mathematical verification

$$\text{var}(aX) = E[(aX)^2] - (E[aX])^2$$

$$\text{var}(aX) = a^2E[X^2] - a^2(E[X])^2$$

$$\text{var}(aX) = a^2 (E[X^2] - (E[X])^2)$$

$$\text{var}(aX) = a^2\text{var}(X)$$

2. Write an R program to exemplify the following property of the variance

For X_1, \dots, X_n independent

$$\text{var} \left(\sum_{i=1}^n X_i \right) = \sum_{i=1}^n \text{var}(X_i) \quad (2)$$

using a sequence of $n = 4$ iid random variables $X_i \sim \text{Normal}(0,1)$, using 10000 repetitions.

```
n <- 4
reps <- 10000

## Generate 4 iid N(0,1) random draws 10000 times
Z <- array( rnorm(n*reps), c(n, reps) )

## Find the variances of the 4 iid random draws
fourVariances <- apply(Z, 1, var)

## Verify that the variance of the sum is roughly the sum of the variances
varOfSum <- var( apply(Z, 2, sum) )

print(varOfSum)
print(sum(fourVariances))
```

Here is a mathematical verification

$$\begin{aligned} \text{var}(X_1 + X_2) &= E[(X_1 + X_2)^2] - (E[X_1 + X_2])^2 \\ &= E[(X_1^2 + 2X_1X_2 + X_2^2)] - (E[X_1] + E[X_2])^2 \\ &= E[X_1^2] + 2E[X_1X_2] + E[X_2^2] - (E[X_1])^2 + (E[X_2])^2 \end{aligned}$$

For X_1 and X_2 independent, $E[X_1X_2] = E[X_1]E[X_2]$.

$$\begin{aligned} \text{var}(X_1 + X_2) &= E[X_1^2] + 2E[X_1]E[X_2] + E[X_2^2] - (E[X_1])^2 - 2E[X_1]E[X_2] - (E[X_2])^2 \\ &= E[X_1^2] - (E[X_1])^2 + E[X_2^2] - (E[X_2])^2 \\ &= \text{var}(X_1) + \text{var}(X_2) \end{aligned}$$

Continuing on by setting $X_1 = X_1 + X_2$, and $X_2 = X_3 \dots$

$$\text{var} \left(\sum_{i=1}^n X_i \right) = \sum_{i=1}^n \text{var}(X_i)$$

3. An Exponential random variable X is a continuous random variable that takes on values between 0 and ∞ . The probability density function for X is

$$q(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

Where λ is the parameter of the distribution, also known as the failure rate. If X has an Exponential distribution with parameter λ we commonly write $X \sim \text{Exponential}(\lambda)$.

Derive a formula for the expected value and variance of a random variable $X \sim \text{Exponential}(\lambda)$. Since an Exponential random variable is a continuous random variable, you will need to use the continuous analogues of the expected value and variance formulas given in the notes. The expected value for a continuous random variable X is

$$EX = \int xq(x)dx$$

plugging in for our Exponential case:

$$EX = \int_0^{\infty} xq(x)dx$$

$$EX = \int_0^{\infty} x\lambda e^{-\lambda x} dx$$

$$EX = \lambda \left[-x \frac{1}{\lambda} e^{-\lambda x} - \frac{1}{\lambda^2} e^{-\lambda x} \right]_0^{\infty}$$

$$EX = \frac{1}{\lambda}$$

The Variance of a continuous random variable X is

$$\text{var}(X) = \int (x - EX)^2 q(x) dx$$

$$\text{var}(X) = E(X^2) - (EX)^2$$

plugging in for our Exponential case:

$$\text{var}(X) = \int_0^{\infty} x^2 \lambda e^{-\lambda x} dx - \frac{1}{\lambda^2}$$

$$= \lambda \left[-x^2 \frac{1}{\lambda} e^{-\lambda x} - 2x \frac{1}{\lambda^2} e^{-\lambda x} - 2 \frac{1}{\lambda^3} e^{-\lambda x} \right]_0^{\infty} - \frac{1}{\lambda^2}$$

$$= \frac{2}{\lambda^2} - \frac{1}{\lambda^2}$$

$$\text{var} X = \frac{1}{\lambda^2}$$

4. Derive a formula for the variance of the sample mean \bar{X} , of n iid Exponential random variables with failure rate λ .

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

The sample mean is a random variable with an associated sampling distribution. We can take advantage of two properties that we exemplified in questions 1 and 2.

$$\text{var}(\bar{X}) = \text{var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n \text{var}(X_i)$$

Since our X_1, \dots, X_n are iid Exponential(λ) we have already calculated the variance of a single X_i to be $\frac{1}{\lambda^2}$, plugging this into the above equation,

$$\text{var}(\bar{X}) = \frac{1}{n^2} \sum_{i=1}^n \frac{1}{\lambda^2} = \frac{1}{n^2} \frac{n}{\lambda^2}$$

$$\text{var}(\bar{X}) = \frac{1}{n\lambda^2}$$

5. Write a simulation program in R in which the variance of the sample mean of n iid Exponential random variables with failure rate λ is estimated. Your program should consider all combinations of sample sizes $n = 10, 20, 40, 80$ and failure rates $\lambda = 1, 2, 3, 4$. To simulate a vector of n iid Exponential random variables with failure rate λ in R use

```
n <- 10
nExponentials <- rexp(n, lambda)
```

Your final result should be a table in which the value calculated in the simulation appears next to the theoretical value based on the result from part (a).

Solution:

```
V <- NULL
reps <- 10000

for (lambda in c(1,2,3,4))
{
  for (n in c(10,20,40,80))
  {
    X <- array(rexp(n*reps, lambda), c(n,reps))
    M <- apply(X, 2, mean)
    V <- rbind(V, c(lambda, n, var(M), 1/(n*lambda^2)))
  }
}
```

6. Derive the EX and $\text{var}(X)$ for $X \sim \text{Poisson}(\lambda)$. Note that X is a discrete random variable, thus we will use the discrete versions of the formulas. The probability mass function for X is

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad \text{for } x = 0, 1, 2, \dots$$

The discrete version of EX is

$$EX = \sum_x x \cdot P(X = x)$$

$$EX = \sum_{x=0}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x!}$$

$$EX = 0 + \sum_{x=1}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x!}$$

$$\frac{x}{x!} = \frac{x}{x(x-1)\dots} = \frac{1}{(x-1)!}$$

let $k = x - 1$, so replace x with $k + 1$

$$EX = \sum_{k=0}^{\infty} \frac{\lambda^{k+1} e^{-\lambda}}{k!}$$

$$EX = \lambda \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!}$$

$$EX = \lambda \sum_{k=0}^{\infty} P(X = k)$$

$$EX = \lambda$$

In order to derive the variance, we can find the $E(X^2)$

$$E(X^2) = \sum_{x=0}^{\infty} x^2 \frac{\lambda^x e^{-\lambda}}{x!}$$

$$E(X^2) = 0 + \sum_{x=1}^{\infty} x^2 \frac{\lambda^x e^{-\lambda}}{x!}$$

$$\frac{x^2}{x!} = \frac{xx}{x(x-1)\dots} = \frac{x}{(x-1)!}$$

$$E(X^2) = 0 + \sum_{x=1}^{\infty} \frac{x \lambda^x e^{-\lambda}}{(x-1)!}$$

let $k = x - 1$, so replace x with $k + 1$

$$\begin{aligned}
E(X^2) &= \sum_{k=0}^{\infty} \frac{(k+1)\lambda^{k+1}e^{-\lambda}}{(k)!} \\
E(X^2) &= \sum_{k=0}^{\infty} \frac{k\lambda^{k+1}e^{-\lambda}}{(k)!} + \sum_{k=0}^{\infty} \frac{\lambda^{k+1}e^{-\lambda}}{(k)!} \\
E(X^2) &= \lambda \sum_{k=0}^{\infty} \frac{k\lambda^k e^{-\lambda}}{(k)!} + \lambda \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{(k)!} \\
E(X^2) &= \lambda EX + \lambda \\
E(X^2) &= \lambda^2 + \lambda \\
\text{var}(X) &= E(X^2) - (EX)^2 \\
\text{var}(X) &= \lambda^2 + \lambda - \lambda^2 \\
\text{var}(X) &= \lambda
\end{aligned}$$

We have shown that the $EX = \text{var}(X) = \lambda$.

7. Derive the formula for $\text{var}(\bar{X})$, where \bar{X} is the sample mean of n iid Poisson random variables with mean λ .

Using the formula we derived earlier,

$$\begin{aligned}
\text{var}(\bar{X}) &= \text{var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n \text{var}(X_i) \\
\text{var}(\bar{X}) &= \frac{1}{n^2} n\lambda \\
\text{var}(\bar{X}) &= \frac{\lambda}{n}
\end{aligned}$$

8. Perform the same simulation in R as we did in question 5.

Solution:

```
V <- NULL
reps <- 10000

for (lambda in c(1,2,3,4))
{
  for (n in c(10,20,40,80))
  {
    X <- array(rpois(n*reps, lambda), c(n,reps))
    M <- apply(X, 2, mean)
    V <- rbind(V, c(lambda, n, var(M), lambda/n))
  }
}
```