S&DS 238 Problem Set #5

Shaun Radgowski

October 16, 2020

Problem 1

(1a)

This is to say that $P\{T > x\} = e^{-\lambda x},\ P\{U > x\} = e^{-\mu x}$.

$$P\{S > x\} = P\{\min(T, U) > x\} = P\{T > x,\ U > x\} = P\{T > x\}\ P\{U > x\}$$

$$= e^{-\lambda x} e^{-\mu x} = e^{-(\lambda+\mu)x}$$

This is true because T and U are independent. This shows (through parallelism) that S is distributed as $\text{Exp}(\lambda + \mu)$ for $x > 0$, and 0 elsewhere.

(1b)

Because T and U are independent:

$$F_{T,U}(t, u) = P\{T \leq t,\ U \leq u\} = P\{T \leq t\}P\{U \leq u\} = (1 - e^{-\lambda t})(1 - e^{-\mu u})$$

$$f_{T,U}(t, u) = \tfrac{\partial^2 F_{T,U}(t,u)}{\partial t \partial u} = \tfrac{\partial}{\partial t}\tfrac{\partial}{\partial u}(1 - e^{-\lambda t} - e^{-\mu u} + e^{-\lambda t - \mu u}) = \lambda\mu e^{-\lambda t - \mu u}$$

$$P\{T < U\} = \textstyle\int_0^\infty \int_0^u f_{T,U}(t, u)\ dt\ du = \lambda\mu \int_0^\infty \int_0^u e^{-\lambda t - \mu u}\ dt\ du$$

$$= \lambda\mu \textstyle\int_0^\infty -\tfrac{1}{\lambda} e^{-\lambda t - \mu u}|_{t=0}^{t=u}\ du = \lambda\mu \int_0^\infty -\tfrac{1}{\lambda}(e^{-(\lambda+\mu)u} - e^{-\mu u})\ du$$

$$= -\mu(\tfrac{e^{-u(\lambda+\mu)}}{-(\lambda+\mu)} - \tfrac{e^{-\mu u}}{-\mu})|_0^\infty = 0 - \mu(\tfrac{e^0}{\lambda+\mu} - \tfrac{e^0}{\mu}) = -\tfrac{\mu}{\lambda+\mu} + 1 = \tfrac{\lambda}{\lambda+\mu}$$

Problem 2

(2a)

If we substitute $X = W_n$ and $Y = W_{n-1}$ into the Law of Total Expectation, we can show:

$$E(W_n) = \textstyle\sum_{W_{n-1}} E(W_n | W_{n-1} = w_{n-1}) P\{W_{n-1} = w_{n-1}\}$$

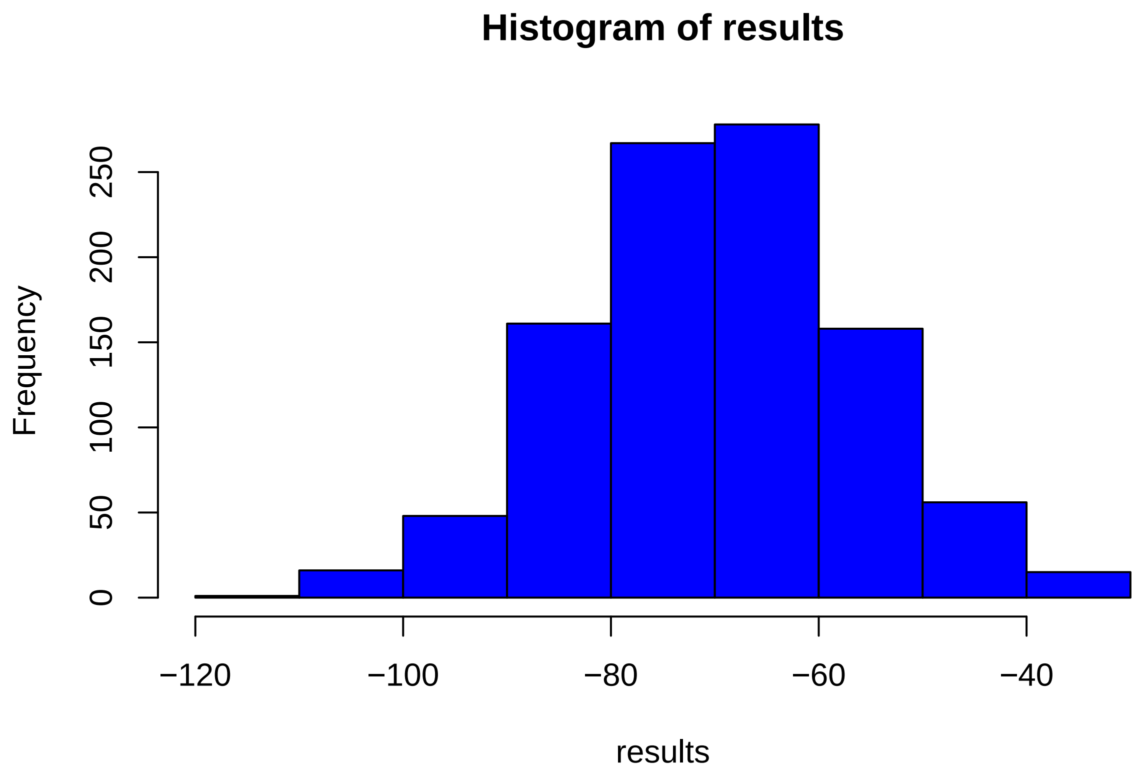
$$= \sum_{W_{n-1}} (0.5(1+b)w_{n-1} + 0.5(1-0.6b)w_{n-1}) \quad P\{W_{n-1} = w_{n-1}\}$$

$$= \sum_{W_{n-1}} (1+0.2b)w_{n-1} \quad P\{W_{n-1} = w_{n-1}\}$$

And because the sum of $w_{n-1}P\{W_{n-1} = w_{n-1}\}$ across all w_{n-1} is equal to $E(W_{n-1})$ by definition, this equation above can be simplified to $E(W_n) = (1+0.2b)E(W_{n-1})$. From this point of view, $b = 1$ appears to yield the greatest expected value of $E(W_n)$ because it maximizes the function for a fixed value of $E(W_{n-1})$. The coefficient in front of b is positive, so a higher b relates to a higher expected value as long as $E(W_{n-1})$ is constant.

(2b)

```r
# Number of simulations
num <- 1000

days <- rep(0, 1460)
days[1] = 1

simulation <- function() {
  for (i in 2:1460) {
    coin <- sample(c(0, 1), size=1, prob=c(0.5, 0.5))

    # Heads
    if (coin == 0) {
      days[i] = 2 * days[i - 1]
    }

    # Tails
    else {
      days[i] = 0.4 * days[i - 1]
    }
  }
  return (log10(days[1460]))
}

results <- rep(0, num)
for (i in 1:num) {
  results[i] <- simulation()
}

hist(results, col="blue")
```



(2c)

The expected value for any given M_i is $0.5(1+b)+0.5(1-0.6b)$, so the expected value for any given $X_i = \log M_i$ is $0.5\log(1+b) + 0.5\log(1-0.6b)$. For $b = 1$, this means that $E(X_n) = 0.5\log(2) + 0.5\log(0.4) \approx -0.05$

The Law of Large Numbers says that $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ approaches the true mean μ as n approaches infinity. We also know that $E(\bar{X}_n) = \mu$. From there, we can say that as n approaches infinity, the sum (representing L_n) approaches negative infinity:

$$\frac{1}{n}\sum_{i=1}^{n} X_i = E(\bar{X}_n), \quad L_n = \sum_{i=1}^{n} X_i = -0.05n \to -\infty$$

(2d)

$$E(\log M_i) = 0.5\log(1+b) + 0.5\log(1-0.6b)$$

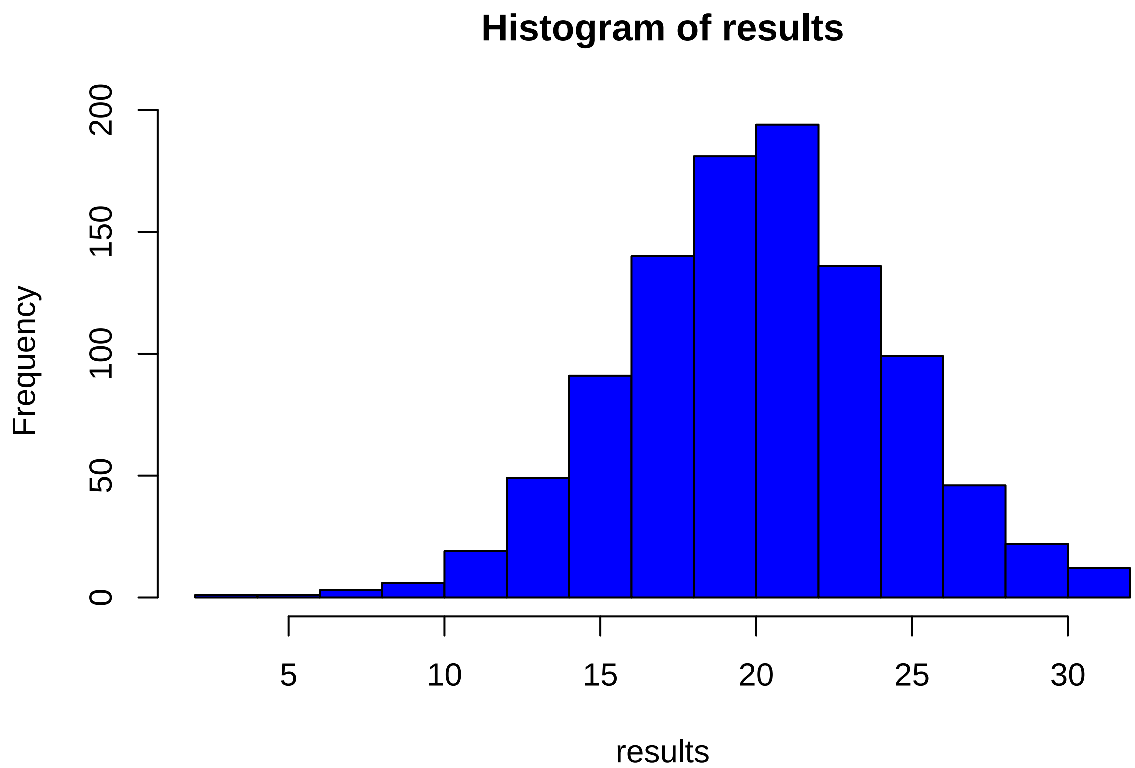
$$\frac{d}{db}E(\log M_i) = \frac{0.5\log_{10} e}{(1+b)} - 0.6\frac{0.5\log_{10} e}{(1-0.6b)} = \frac{0.2171}{(1+b)} - \frac{0.1303}{(1-0.6b)}$$

Setting that expression equal to 0 will find the maximizing value of b:

$$\frac{0.2171}{(1+b)} = \frac{0.1303}{(1-0.6b)}, \ 0.0868 = 0.2606b, \quad b = \frac{0.0869}{0.2606} = 0.33$$

(2e)

```r
# Number of simulations
num <- 1000

days <- rep(0, 1460)
days[1] = 1

simulation <- function() {
  for (i in 2:1460) {
    coin <- sample(c(0, 1), size=1, prob=c(0.5, 0.5))

    # Heads
    if (coin == 0) {
      days[i] = 1.333 * days[i - 1]
    }

    # Tails
    else {
      days[i] = 0.799 * days[i - 1]
    }
  }
  return (log10(days[1460]))
}

results <- rep(0, num)
for (i in 1:num) {
  results[i] <- simulation()
}

hist(results, col="blue")
```




```r
sum(results > 10)/length(results)
```

```
## [1] 0.989
```

Above is the histogram of $\log_{10}$ of the results of playing this game over 4 years with $b = 0.4$. To make at least 10 billion dollars, the result would need to be at least 10 (as $\log_{10} 10,000,000,000 = 10$), which seems to be true for around 99% of the outcomes.

Problem 3

(3a)

In order for Z to also be an unbiased estimator of μ, it would need to be such that $E(Z) = \mu$. From the properties of expectation values:

$$E(Z) = E(aX) + E(bY) = aE(X) + bE(Y) = (a+b)\mu$$

In order for this to be true, it must be true that $a + b = 1$.

(3b)

$$Var(Z) = Var(aX) + Var(bY) + 2(E(aXbY) - E(aX)E(bY))$$

$$= a^2Var(X) + b^2Var(Y) + 2ab(E(XY) - E(X)E(Y))$$

X and Y are independent, so their covariance is 0, thus eliminating the final term in the expression above.

$$= a^2\sigma_X^2 + b^2\sigma_Y^2$$

If we treat the variances σ_X^2 and σ_Y^2 as constants, then to minimize the above expression given the condition that $a + b = 1$, we could use a Lagrangian multiplier and set the gradient equal to 0.

$$\mathcal{L}(a, b, \lambda) = a^2\sigma_X^2 + b^2\sigma_Y^2 - \lambda(a + b - 1)$$

$$\frac{\partial\mathcal{L}(a,b,\lambda)}{\partial a} = 2a\sigma_X^2 - \lambda = 0, \quad a = \frac{\lambda}{2\sigma_X^2}$$

$$\frac{\partial\mathcal{L}(a,b,\lambda)}{\partial b} = 2b\sigma_Y^2 - \lambda = 0, \quad b = \frac{\lambda}{2\sigma_Y^2}$$

$$\frac{\partial\mathcal{L}(a,b,\lambda)}{\partial \lambda} = a + b - 1 = 0, \quad a + b = 1$$

$$\lambda = 2\sigma_X^2 a = 2\sigma_Y^2 b, \quad a + \frac{\sigma_X^2 a}{\sigma_Y^2} = 1, \quad a = \frac{1}{1+\frac{\sigma_X^2}{\sigma_Y^2}}, \quad b = \frac{1}{1+\frac{\sigma_Y^2}{\sigma_X^2}}$$

Problem 4

(4a)

Markov's inequality shows that if Y is any nonnegative random variable, then $P\{Y \geq c\} \leq \frac{E(Y)}{c}$. First, we can define another nonnegative random variable X to be $X = (Y - E(Y))$. We can then apply Markov's inequality to X to show that $P\{X \geq c^2\} \leq \frac{E(X)}{c^2}$. From there, we can say:

$$E(X) = E((Y - E(Y))^2) = Var(Y) = \sigma^2$$

$$P\{X \geq c^2\} = P\{(Y - E(Y))^2 \geq c^2\} = P\{|Y - E(Y)| \geq c\}$$

Putting these together with the fact that $E(Y) = \mu$, we can conclude:

$$P\{|Y - \mu| \geq c\} \leq \frac{\sigma^2}{c^2}$$

(4b)

Using Chebyshev's inequality, we know that for large n:

$$P\{|\bar{X}_n - \mu| \geq c\} \leq \frac{\sigma^2}{nc^2}$$

In this instance, we can determine that c is 0.5 and thus say:

$$P\{|\bar{X}_n - 7| \leq c\} \geq 1 - \frac{4}{nc^2}, \quad P\{7 - c \leq \bar{X}_n \leq 7 + c\} = P\{6.5 \leq \bar{X}_n \leq 7.5\} \geq 1 - \frac{16}{n}$$

Thus, in order for the right side of the inequality to be at least 0.9, n must be the following (rounding to the nearest integer):

$$1 - \frac{16}{n} \geq 0.9, \quad 0.1 \geq \frac{16}{n}, \quad n \geq \frac{16}{0.1} \approx 160$$

Problem 5

(5a)

(i)

Through the Taylor series expansion $e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}$, we can see that for the sum over values $k = 0, 1, ...$ the expression $\frac{\lambda^k}{k!}$ is equivalent to e^{λ}. Thus, summed over all nonnegative values of k, the expression in question becomes:

$$e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}e^{\lambda} = e^0 = 1$$

This is a valid probability mass function for these values of k because the total probability across all values of k sums to 1, and the probability is nonnegative for all nonnegative values of k (as expected for a PMF).

(ii)

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

We can eliminate the term where $k = 0$ (as the entire term would thus be 0), and consolidate the k with the $k!$.

$$= \sum_{k=1}^{\infty} e^{-\lambda} \frac{\lambda^k}{(k-1)!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}$$

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

(5b)

If we define $g(X) = (-2)^X$, then the LOTUS states:

$$E(g(X)) = \sum g(k) f_X(k) = \sum_{k=0}^{\infty} (-2)^k e^{-\lambda} \frac{\lambda^k}{k!} = \sum_{k=0}^{\infty} e^{-\lambda} \frac{(-2\lambda)^k}{k!}$$

Bringing back the Taylor series for the exponential, this time with $z = -2\lambda$:

$$E((-2)^X) = e^{-\lambda} e^{-2\lambda} = e^{-3\lambda} = \theta$$

(5c)

"Unbiased" means that the expected value of the estimator is the thing that we are trying to measure. In this case, claiming that $\delta(X)$ is an unbiased of estimator of θ is equivalent to claiming that the expectation value of $\delta(X)$ is θ. We know this to be true, as we proved it in part B: $E((-2)^X) = \theta$.