Math 318 Homework 5

Xander Naumenko

02/03/23

Question 1a. The meaning of $1 - F_M(x)$ is the probability that the minimum is above a given value x. This occurs when each of the X_i are above x, which happens with probability each $1 - x$. Thus we have

$$1 - F_M(x) = (1-x)^n \implies F_M(x) = 1 - (1-x)^n.$$

Question 1b. The PDF is just the derivative of the CDF:

$$f_M(x) = \frac{d}{dx} F_M(x) = n(1-x)^{n-1}.$$

Question 1c. To find the mean we can integrate:

$$\mu = \int_0^1 nx(1-x)^{n-1}dx = x(1-x)^n\Big|_0^1 + \int_0^1 (1-x)^n dx = \frac{1}{n+1}.$$

Similarly for the variance, omitting the boundary terms since as in the previous calculation they obviously go to zero:

$$\mathrm{Var}(M) = \int_0^1 nx^2(1-x)^{n-1}dx = \int_0^1 2x(1-x)^n dx = \int_0^1 = \int_0^1 \frac{2}{n+1}(1-x)^{n+1}dx = \frac{2}{(n+1)(n+2)}.$$

Question 1d. Let $y = xn$. Then the CDF of Y is

$$F_Y(y) = F_M(\frac{y}{n}) = 1 - (1 - \frac{y}{n})^n.$$

$$f_Y(y) = \frac{d}{dy} F_Y(y) = (1 - \frac{y}{n})^{n-1}.$$

Question 1e. Taking the limit and using a definition of e stated in class:

$$\lim_{n\to\infty} = (1 - \frac{y}{n})(1 - \frac{y}{n})^n = \lim_{n\to\infty} (1 - \frac{y}{n})e^{-y} = e^{-y}.$$

Question 2. Clearly $E[X] = 0.5$ and $E[Y] = -0.5$ since they're uniform distributions. We also have

$$E[X^2] = E[Y^2] = \int_0^1 x^2 dx = \frac{1}{3}.$$

Now computing the covariance using the definition

$$\mathrm{cov}(X, X^2) = E\left[(X - E[X])(X^2 - E[X^2])\right] = E[XX^2] - E[X]E[X^2]$$

$$= \int_0^1 xxx^2 dx - \frac{1}{6} = \frac{1}{5} - \frac{1}{6} = \frac{1}{30}.$$

Doing the same calculation for Y:

$$\mathrm{cov}(Y, Y^2) = E\left[(Y - E[Y])(Y^2 - E[Y^2])\right] = E[YY^2] - E[Y]E[Y^2].$$

$$= \int_{-1}^0 yyy^2 dy - \frac{1}{6} = -\frac{1}{5} - \frac{1}{6} = -\frac{11}{30}.$$

Question 3a.

Question 4a. Integrating to find the marginal distribution:

$$f(x) = \int_0^\infty \lambda^2 e^{-\lambda z} 1_{0 \le x \le z} dz = \int_x^\infty \lambda^2 e^{-\lambda z} dz = \lambda e^{-\lambda x}.$$

Question 4b. Doing same as previous question:

$$f(z) = \int_0^\infty \lambda^2 e^{-\lambda z} 1_{0 \le x \le z} dx = \int_0^z \lambda^2 e^{-\lambda z} dx = \lambda^2 z e^{-\lambda z}.$$

Question 4c. As the hint suggests, we can take the derivative of the joint cumulative distribution function of X and Y. To find what this is, we can first find the joint cumulative distribution of X, Z:

$$F_{X,Z}(x, z) = \int_0^z \int_0^x \lambda^2 e^{-\lambda z'} 1_{0 \le x' \le z'} dx' dz' = \lambda \left(1 - e^{-\lambda z}\right) \min(x, z).$$

We can then do the calculation to find the joint density

$$f_{X,Y}(x, y) = \frac{d}{dx}\frac{d}{dy} P(X \le x, Y \le y) = \frac{d}{dx}\frac{d}{dy} P(X \le x, Z \le x + y) = \frac{d}{dx}\frac{d}{dy} F_{X,Z}(x, x + y).$$

$$= \frac{d}{dx}\frac{d}{dy}\left(\lambda\left(1 - e^{-\lambda(x+y)}\right)\min(x, x+y)\right) = \frac{d}{dx}\frac{d}{dy}\lambda x\left(1 - e^{-\lambda(x+y)}\right).$$

$$= \frac{d}{dx}\lambda^2 x e^{-\lambda(x+y)} = \lambda^2 e^{-\lambda(x+y)} - \lambda^3 x e^{-\lambda(x+y)} = (1 - \lambda x)\lambda^2 e^{-\lambda(x+y)}.$$

Question 5a. The following python code was used to produce the output. The result was $P(A) \approx 0.0168$.

```python
from math import comb
n = 100
p = 0.6
P = sum([comb(n, k) * p**k * (1-p)**(n-k) for k in range(0, 50)])
print(P)
```

Question 5b. Let $X_1, X_2, \ldots, X_{100}$ be a set of independent Bernoulli variables with parameter $p = 0.6$, so $X = \sum_{i=1}^{100} X_i$. The mean of X is clearly $np = 60$ and the variance of each X_i is $0.6 \cdot 0.4 = 0.24$, so by the central limit theorem we have

$$Z = \frac{X - 60}{\sqrt{24}} \approx N(0, 1).$$

Question 5c. Using the fact that $X = \sqrt{24}Z + 60$, we get:

$$P(X \leq 50) = P(\sqrt{24}Z + 60 \leq 50) = P(Z \leq -\frac{10}{\sqrt{24}}) \approx \Phi(-\frac{10}{\sqrt{24}}) \approx 0.0206.$$

Question 5d. Using the exact same approach as part c except with 51 instead of 50:

$$P(X < 51) = P(\sqrt{24}Z + 60 < 51) = P(Z < -\frac{9}{\sqrt{24}}) \approx \Phi(-\frac{9}{\sqrt{24}}) \approx 0.0206.$$

Question 6a. We can compute the probability distribution of each S_n by taking convolutions of the pdfs of S_{n-1} X_n. The following code does this, the resultant graphs can be seen in figure 1

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import norm

ns = (1,2,3,4,5,10,50)

u = np.array([1/3, 1/3, 1/3])
S = [u]

for i in range(1, max(ns)+1):
    s = np.convolve(u, S[-1])
    S.append(s)

fig,ax = plt.subplots(nrows=4, ncols=2, sharey=True)
for i,n in enumerate(ns):
    a = ax[i//2,i%2]
    a.plot(S[n-1], label='$S_n$')
    a.title.set_text(f'n={n}')
    mu = n
    var = 2/3 * (n)

    x = np.arange(len(S[n-1]))
    a.plot(norm.pdf(x, mu, var**0.5), label=f'$N({mu},{round(var,1)})$')
    a.legend()

plt.show()
```

Question 6b.

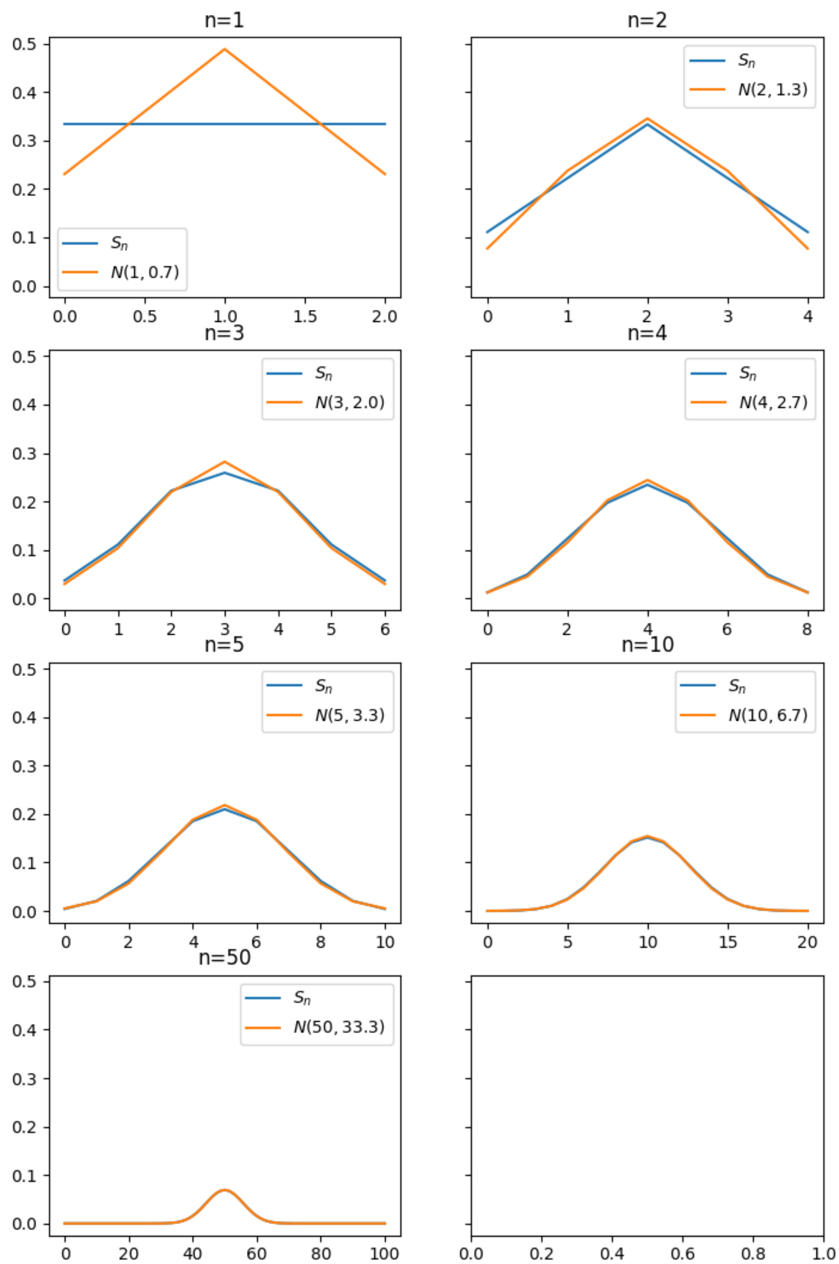

Figure 1: Graphs for question 6a