HOMEWORK 9

PSTAT 120A: Summer 2022

Due: 11:59pm on Tuesday, July 26, 2022
Instructor: Ethan P. Marzban

Instructions:

- Please submit your work to Gradescope by no later than the due date posted above.
- Sometime on Tuesday morning, the solutions to this homework will become available (yes, before the due date).
- Be sure to show your work; correct answers with no supporting work will not be awarded full points.
- 2 randomly selected questions/parts will be graded, but you must still turn in your work for all problems in order to be eligible to earn full credit.

1. The hit new restaurant *Gaucho Gourmande* is trying to set their menu prices.

   It is known that the true cost of making their signature dish varies from day to day (due to changes in market pricing of ingredients). We assume that the cost of manufacturing this dish on any given day is a random variable with mean \$30 and standard deviation \$2, independent of all other days.

   The owners of *Gaucho Gourmande* would like to set the price of the dish to be a fixed number \$$m$, and would like to ensure that with probability 0.99 the revenue generated from selling the dish exceeds the true cost of manufacturing it in a 100-day period. [Assume, somewhat unrealistically, that the restaurant sells only one dish per day.]

   (a) Find the value of m using Chebyshev's Inequality
   (b) Find the value of m using the Central Limit Theorem

Solution: Let X_i denote the cost of manufacturing the dish on day i: then $\mathbb{E}[X_i] = 30$ and $\mathrm{SD}(X) = 2$. Additionally, if C denotes the total cost of manufacturing this dish for 100 days,

$$C = \sum_{i=1}^{100} X_i$$

and $\mathbb{E}[C] = 100 \cdot 30 = 3{,}000$ with $\mathrm{Var}(C) = 100 \cdot (2^2) = 400$.

On the other hand, the revenue generated by selling the dish for \$100 days is $100m$; therefore, we seek m such that

$$\mathbb{P}(C \leq 100m) \geq 0.99$$

(a) By Chebyshev's,

$$\begin{aligned}\mathbb{P}(C \geq 100m) &= \mathbb{P}(C - 3000 \geq 100m - 3000) \\ &\leq \mathbb{P}(|C - 3000| \geq 100m - 3000) \\ &\leq \frac{400}{(100m - 3000)^2}\end{aligned}$$

We need this probability to be at most $1 - 0.99 = 0.01$ which leads to

$$\frac{400}{(100m-3000)^2} \leq 0.01 \implies m \geq \frac{1}{100}\left(\sqrt{\frac{400}{0.01}} + 3000\right) = \$32.00$$

(b) By the CLT,

$$C \overset{d}{\approx} \mathcal{N}\left(3000,\ 400\right)$$

and so

$$\mathbb{P}(C \leq 100m) = \mathbb{P}\left(\frac{C - 3000}{20} \leq \frac{100m - 3000}{20}\right)$$

$$\Phi\left(\frac{100m - 3000}{20}\right) \overset{!}{\geq} 0.99$$

From the $z-$table, we can see that $\Phi(2.33) \approx 0.9900$, so

$$\frac{100m - 3000}{20} = 2.33 \implies m = \$30.47$$

2. After Session A ends, Ethan is a bit low on funds. As such, he decides to go back to basics and start up a lemonade stand! He's not very good, though (customer reviews seem to be rife with complaints about how he manages to somehow turn everyday scenarios into math problems?); on average, he only makes about \$10 a day. His true earnings on any given day is, however, a random variable. Additionally, Ethan finds that his earnings are correlated across days; intuitively, it makes sense that earnings on days that are farther apart would be less correlated, so the following covariance structure is adopted:

$$\text{Cov}(X_i, X_j) = \left(\frac{1}{2}\right)^k \text{ for } |i - j| = k, \quad i, j \in \{1, 2, \cdots\}$$

Ethan is interested in his net (total) earnings over a 30-day period.

(a) What is the expected total earnings Ethan will make over these 30 days?

Solution: Let X_i denote the earnings on a given day, and let T denote the total earnings over a 30-day period. Then, $\mathbb{E}[X_i] = 10$ (given in the problem) and $T = \sum_{i=1}^{30} X_i$, meaning

$$\mathbb{E}[T] = \mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{30} \mathbb{E}[X_i] = 30 \cdot \mathbb{E}[X_1] = 300$$

(b) What is the variance of the total earnings Ethan will make over these 30 days?

Solution: In general, we recall that

$$\text{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \text{Var}(X_i) + 2\sum_{i<j} \text{Cov}(X_i, X_j)$$

It may be useful to write out a few terms in the variance-covariance matrix:

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & 1/2 & 1/4 & 1/8 & 1/16 & \cdots \\ 1/2 & 1 & 1/2 & 1/4 & 1/8 & \cdots \\ 1/4 & 1/2 & 1 & 1/2 & 1/4 & \cdots \\ 1/8 & 1/4 & 1/2 & 1 & 1/2 & \cdots \\ 1/16 & 1/8 & 1/4 & 1/2 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \end{pmatrix}$$

The $\sum_{i<j} \mathrm{Cov}(X_i, X_j)$ quantity amounts to summing up all of the values in the lower-triangular portion of $\mathbf{\Sigma}$; hence,

$$\begin{aligned}\sum_{i<j} \mathrm{Cov}(X_i, X_j) &= \sum_{k=2}^{30} \sum_{j=1}^{k-1} \left(\frac{1}{2}\right)^j \\ &= \sum_{k=2}^{30} \frac{\left(\frac{1}{2}\right) - \left(\frac{1}{2}\right)^k}{1 - \left(\frac{1}{2}\right)} = \sum_{k=2}^{30} \left[1 - \left(\frac{1}{2}\right)^{k-1}\right] \\ &= 29 - \frac{\left(\frac{1}{2}\right) - \left(\frac{1}{2}\right)^{30}}{1 - \left(\frac{1}{2}\right)} = 29 - 1 + \left(\frac{1}{2}\right)^{29} = 28 + \left(\frac{1}{2}\right)^{29}\end{aligned}$$

Therefore,

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{30} (1) + 2 \cdot \left[28 + \left(\frac{1}{2}\right)^{29}\right] = 30 + 56 + \left(\frac{1}{2}\right)^{28} = 86 + \left(\frac{1}{2}\right)^{28} \approx 86$$

(c) Approximate the probability that Ethan will earn more than \$310 over these 30 days.

Solution: By the CLT, the distribution of S will be roughly normal. Since we computed the mean and variance of S above, we can conclude

$$S \overset{\mathrm{d}}{\approx} \mathcal{N}\left(300,\ 86 + \left(\frac{1}{2}\right)^{28}\right)$$

We seek $\mathbb{P}(S > 310)$; we compute this by standardizing

$$\mathbb{P}(S > 310) = \mathbb{P}\left(\frac{S - 300}{\sqrt{86 + \left(\frac{1}{2}\right)^{28}}} > \frac{310 - 300}{\sqrt{86 + \left(\frac{1}{2}\right)^{28}}}\right) \approx 1 - \Phi\left(\frac{10}{\sqrt{86 + \left(\frac{1}{2}\right)^{28}}}\right) \approx 1 - \Phi(1.08)$$

which, if we wanted a numerical value, equates to approximately 14%

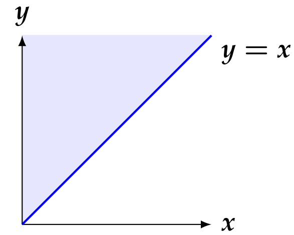
3. Let (X, Y) be a pair of bivariate random variables with joint probability density function (p.d.f.) given by

$$f_{X,Y}(x, y) = \begin{cases} \dfrac{\lambda}{y} \cdot e^{-\lambda y} & \text{if } 0 < x < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

where $\lambda > 0$ is a fixed, deterministic constant.

(a) Show that $f_{X,Y}(x, y)$ is a valid probability density function.

Solution: Clearly $f_{X,Y}(x, y) \geq 0$, since the only portion over which the density is nonzero lies within the first quadrant meaning $y \geq 0$, and $e^{-\lambda y} \geq 0$ for all real y. Additionally, the support looks like:



Therefore, we see that

$$\begin{aligned}\iint_{\mathbb{R}^2} f_{X,Y}(x,y)\,\mathrm{d}A &= \int_0^\infty \int_0^y \frac{\lambda}{y} e^{-\lambda y}\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_0^\infty \cancel{y} \cdot \frac{\lambda}{\cancel{y}} e^{-\lambda y}\,\mathrm{d}y = \int_0^\infty \lambda e^{-\lambda y}\,\mathrm{d}y = 1 \checkmark\end{aligned}$$

Hence, since $f_{X,Y}(x,y) \geq 0$ for all $(x,y) \in \mathbb{R}^2$ and $f_{X,Y}(x,y)$ integrates to unity, $f_{X,Y}(x,y)$ is a valid probability density function. (An aside: it is impossible to integrate in the order $\mathrm{d}y\,\mathrm{d}x$, as e^{-y}/y is actually another function that does not have an elementary antiderviative.)

(b) Identify the marginal distribution of Y.

Solution: Let's first find the marginal p.d.f.:

$$\begin{aligned}f_Y(y) &= \int_{-\infty}^\infty f_{X,Y}(x,y)\,\mathrm{d}x \\ &= \int_0^y \frac{\lambda}{y} e^{-\lambda y}\,\mathrm{d}x \\ &= \lambda e^{-\lambda y} \cdot \frac{1}{y} \cdot \int_0^y \mathrm{d}x = \lambda e^{-\lambda y}\end{aligned}$$

which, since $S_Y = [0, \infty)$, shows that $Y \sim \mathrm{Exp}(\lambda)$

(c) Identify the conditional distribution of $(X \mid Y)$.

Solution: We first find the conditional density $f_{X|Y}(x \mid y)$:

$$\begin{aligned}f_{X|Y}(x \mid y) &= \frac{f_{X,Y}(x,y)}{f_Y(y)} \\ &= \frac{\frac{\lambda}{y} e^{-\lambda y} \cdot \mathbb{1}_{\{x \in [0,y]\}} \cdot \mathbb{1}_{\{y \geq 0\}}}{\lambda e^{-\lambda y} \cdot \mathbb{1}_{\{y \geq 0\}}} = \frac{1}{y} \cdot \mathbb{1}_{\{x \in [0,y]\}}\end{aligned}$$

which shows that $(X \mid Y = y) \sim \mathrm{Unif}(0, y)$ and so

$$(X \mid Y) \sim \mathrm{Unif}[0, Y]$$

(d) Using your answer to the above parts of this question, and **without performing any integration**, compute $\mathbb{E}[X]$.

Solution: The Law of Iterated Expectations tells us that $\mathbb{E}[X] = \mathbb{E}\left[\mathbb{E}[X \mid Y]\right]$. From part (c) we conclude

$$\mathbb{E}[X \mid Y] = \frac{Y + 0}{2} = \frac{Y}{2}$$

Additionally, by part (b), $\mathbb{E}[Y] = 1/\lambda$ meaning

$$\mathbb{E}[X] = \mathbb{E}\left[\frac{Y}{2}\right] = \frac{1}{2} \cdot \frac{1}{\lambda} = \boxed{\frac{1}{2\lambda}}$$