18.650. Fundamentals of Statistics
Spring 2026. Problem Set 1

Due Wednesday, Feb 18

Problem 1 (Probability Review)**.** *Please use the accompanying bubble sheet for submitting your solutions for this problem.*

In what follows, Φ is the CDF of the standard Gaussian (Normal) distribution.

Let X be a random variable taking values between 0 and π, with pdf given by

$$f(x) = c \sin x, \qquad x \in [0, \pi].$$

1. What is the value of c? (a) π (b) $\mathbf{1/2}$ (c) 2 (d) $1/\pi$
2. What is $\mathbb{E}[X]$? (a) $\boldsymbol{\pi/2}$ (b) π (c) 1 (d) $1/2$

Let X be a Gaussian random variable with mean $\mu > 0$ and variance σ^2.

3. What is $\mathbb{E}[X]$? (a) 0 (b) σ^2 (c) $\mu^2 + \sigma^2$ (d) $\boldsymbol{\mu}$
4. What is $\mathbb{V}[X]$? (a) $\boldsymbol{\sigma^2}$ (b) $\mu^2 + \sigma^2$ (c) μ^2 (d) σ
5. What is $\mathbb{E}[X^2]$? (a) μ^2 (b) $(\mu + \sigma)^2$ (c) $\boldsymbol{\mu^2 + \sigma^2}$ (d) σ^2
6. What is $\mathbb{E}[X^3]$? (a) $\mu^3 + 3\mu\sigma$ (b) $\boldsymbol{\mu^3 + 3\mu\sigma^2}$ (c) μ^3 (d) $3\mu\sigma^2$
7. What is $\mathbb{V}[X^2]$? (a) $4\mu^2\sigma^2$ (b) $(\mu^2 + \sigma^2)^2$ (c) $2\sigma^4$ (d) $\boldsymbol{4\mu^2\sigma^2 + 2\sigma^4}$
8. What is $\mathbb{P}(X > 0)$ in terms of Φ?
   (a) $\boldsymbol{\Phi(\mu/\sigma)}$ (b) $\Phi(\mu/\sigma^2)$ (c) $1 - \Phi(\mu/\sigma)$ (d) $\Phi(\mu)$

Let $X \sim \text{Lognormal}(\mu, \sigma^2)$, i.e., $\log X \sim \mathcal{N}(\mu, \sigma^2)$ with $\sigma > 0$.

9. What is $\mathbb{E}[X]$? (a) e^{μ} (b) $\boldsymbol{e^{\mu + \sigma^2/2}}$ (c) $e^{\mu + \sigma^2}$ (d) $\mu + \sigma^2/2$
10. What is the median of X? (a) $e^{\mu + \sigma^2/2}$ (b) $e^{\mu - \sigma^2}$ (c) $\boldsymbol{e^{\mu}}$ (d) μ
11. What is $\mathbb{P}(X > 1)$ in terms of Φ?
    (a) $1 - \Phi(\mu/\sigma)$ (b) $\Phi(\mu)$ (c) $\Phi(\mu/\sigma^2)$ (d) $\boldsymbol{\Phi(\mu/\sigma)}$

Let $X, Y \sim \text{Lognormal}(0, 1)$ be independent.

12. What is $\mathbb{P}(XY > 1)$? (a) $1/4$ (b) $\mathbf{1/2}$ (c) $1/\sqrt{2\pi}$ (d) $1/e$

Let X be a random variable such that

$$X = \begin{cases} 1 & \text{with probability } p \\ -1 & \text{with probability } 1-p \end{cases}$$

for some $p \in [0, 1]$.

13. What is $\mathbb{E}[X]$?

    (a) $-p$ (b) p (c) $1-2p$ (d) $\boldsymbol{2p-1}$

14. What is $\mathbb{V}[X]$?

    (a) $p(1-p)$ (b) $4p-p^2$ (c) $\boldsymbol{4p(1-p)}$ (d) $4p^2(1-p)$

15. For what p is $\mathbb{V}[X]$ maximized?

    (a) 1 (b) 0 (c) **0.5** (d) $1/\sqrt{2}$

16. What is $\mathbb{E}[X^k]$?

    (a) p^k (b) $p^k - (1-p)^k$ (c) $p(-1)^k + (1-p)$ (d) $\boldsymbol{p + (1-p)(-1)^k}$

Let X, Y be two independent standard Gaussian random variables.

17. What is $\mathbb{E}[X^2Y]$?

    (a) **0** (b) 1 (c) 2 (d) 3

18. What is $\mathbb{V}(X+Y)$?

    (a) 0 (b) 1 (c) **2** (d) 3

19. What is $\mathbb{V}(XY)$?

    (a) 0 (b) **1** (c) 2 (d) 3

20. What is $\text{Cov}(X, X+Y)$?

    (a) 0 (b) **1** (c) 2 (d) 3

21. What is $\text{Cov}(X, XY)$?

    (a) **0** (b) 1 (c) 2 (d) 3

Let X be an exponential random variable with parameter $1/2$ that models the lifetime (in years) of a lightbulb.[1]

22. What is (approximately) the probability that the lightbulb will last at least 2 years?

    (a) 0.002 (b) **0.018** (c) 0.180 (d) 0.810

[1] We use the convention from AoS for the parameter of an exponential distribution.

23. Given that the lightbulb has already lasted for at least 3 years, what is (approximately) the probability that it will last for at least two more years?

    (a) 0.002 (b) **0.018** (c) 0.180 (d) 0.810

Let $X_1, \ldots, X_n$ be i.i.d with mean μ and variance σ^2.

24. What is $\mathbb{E}[\sum_{i=1}^n X_i]$? (a) μ (b) $n\sigma$ (c) $\boldsymbol{n\mu}$ (d) σ

25. What is $\mathbb{V}[\sum_{i=1}^n X_i]$? (a) $n^2\sigma^2$ (b) $\boldsymbol{n\sigma^2}$ (c) $n\sigma^2 + n^2\mu^2$ (d) $n\mu$

26. What is $\mathbb{E}\big[\big(\sum_{i=1}^n X_i\big)^2\big]$? (a) $n^2\mu^2$ (b) $n\sigma^2$ (c) $n\mu$ (d) $\boldsymbol{n\sigma^2 + n^2\mu^2}$

27. What is $\mathbb{E}[\frac{1}{n}\sum_{i=1}^n X_i]$? (a) σ (b) $n\sigma^2$ (c) $n\mu$ (d) $\boldsymbol{\mu}$

28. What is $\mathbb{V}[\frac{1}{n}\sum_{i=1}^n X_i]$? (a) μ (b) $\boldsymbol{\sigma^2/n}$ (c) σ^2 (d) $n\mu$

* * *

The following problems are of "show-your-work" type. You get complete points for a solution if you show your complete work. That includes all computations that lead to your answer with appropriate reasoning. Please upload your work to Gradescope.

Problem 2. Let $X_n \sim \mathrm{Unif}\big(-\frac{1}{n}, \frac{1}{n}\big)$ and let X be a random variable such that $\mathbb{P}(X = 0) = 1$.

1. Compute and draw the CDF $F_n(x)$ and $F(x)$ of X_n and X respectively.

   Solution. $F_n(x) = 0$ for $x \leq -1/n$ and $F_n(x) = 1$ for $x \geq 1/n$. For $x \in (-1/n, 1/n)$, $F_n(x)$ is linear, from the value zero at $x = -1/n$ to the value 1 at $x = 1/n$.

   For F, we have $F(x) = 0$ when $x < 0$, and $F(x) = 1$ when $x \geq 0$. (In particular, F is right continuous.)

2. Does $X_n \xrightarrow{\mathbb{P}} X$? (prove or disprove)

   Solution. For any $\epsilon > 0$ we have $\mathbb{P}(|X_n - X| > \epsilon) = \mathbb{P}(|X_n| > \epsilon) = 0$ for all $n \geq 1/\epsilon$. Therefore, $\mathbb{P}(|X_n - X| > \epsilon) \to 0$ as $n \to \infty$, so X_n converges to X in probability.

3. Does $X_n \rightsquigarrow X$? (prove or disprove)

   Solution. X_n converges to X in distribution because it converges in probability.

Problem 3. Let $X \sim \mathcal{N}(2, 1.44)$. Compute the following probabilities:

1. $\mathbb{P}(2X - 1 < 0) = \Phi(-5/4)$

2. $\mathbb{P}\left(\frac{7}{5} \leq X \leq \frac{16}{5}\right) = \Phi(1) - \Phi(-1/2)$

3. $\mathbb{P}\left(X > \frac{16}{5} \mid X > \frac{7}{5}\right) = \frac{1-\Phi(1)}{\Phi(1/2)}$

4. $\mathbb{P}\left(X \le \frac{4}{5} \text{ or } X \ge \frac{16}{5}\right) = 2(1-\Phi(1))$

Problem 4. Let

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\right)$$

Compute the following quantities (show your work):

1. $\mathbb{V}[X] = 1.$

2. $\mathbb{E}[Y^2 + X] = \mathbb{V}(Y) + (\mathbb{E}Y)^2 + \mathbb{E}X = 2 + 0^2 + 1 = 3.$

3. $\mathbb{E}[(X-Y)^2] = \mathbb{E}X^2 + \mathbb{E}Y^2 - 2\mathbb{E}XY = \mathbb{V}(X) + (\mathbb{E}X)^2 + \mathbb{V}(Y) + (\mathbb{E}Y)^2 - 2\operatorname{Cov}(X,Y) - 2(\mathbb{E}X)(\mathbb{E}Y) = 1 + 1^2 + 2 + 0^2 - 2 - 0 = 2.$

4. $\mathbb{V}[X + 2Y] = \mathbb{V}(X) + 4\mathbb{V}(Y) + 4\operatorname{Cov}(X,Y) = 1 + 8 + 4 = 13.$

5. Find $\alpha > 0$ such that $\alpha X = Y$ with probability 1 or prove that no such α exists.

   Solution. Notice that $\mathbb{E}[\alpha X] = \alpha$ while $\mathbb{E}Y = 0$. Since $A = B$ with probability 1 implies $\mathbb{E}A = \mathbb{E}B$, no $\alpha > 0$ can exist with the desired property.

Problem 5. We are testing n lightbulbs. Each bulb independently *passes* some quality check with probability $p \in (0,1)$ and *fails* with probability $1-p$. Let $X_i \in \{0,1\}$ indicate the outcome, where $X_i = 1$ means the bulb passes.

Conditioned on X_i, the lifetime Y_i of bulb i is exponentially distributed:

$$Y_i \mid (X_i = 1) \sim \operatorname{Exp}(\lambda_1), \qquad Y_i \mid (X_i = 0) \sim \operatorname{Exp}(\lambda_0),$$

where $\lambda_0, \lambda_1 > 0$.[2] Assume the pairs (X_i, Y_i) are i.i.d. across i.

1. Let (X, Y) be a copy of (X_1, Y_1) and let $T := XY$. Compute the following quantities in terms of p, λ_0, λ_1:

   $\mathbb{E}[Y]$, $\mathbb{V}(Y)$, $\operatorname{Cov}(X,Y)$, $\mathbb{E}[T]$, $\mathbb{V}(T)$, and $\operatorname{Cov}(X,T)$.

   Solution. We have $\mathbb{E}[X] = p$ and $\operatorname{Var}(X) = p(1-p)$.

   Let $m := \mathbb{E}[Y]$. Then

   $$m = \mathbb{E}[\mathbb{E}[Y \mid X]] = p\lambda_1 + (1-p)\lambda_0.$$

   By the law of total variance,

   $$\begin{aligned}\operatorname{Var}(Y) &= \mathbb{E}[\operatorname{Var}(Y \mid X)] + \operatorname{Var}(\mathbb{E}[Y \mid X]) \\ &= p\lambda_1^2 + (1-p)\lambda_0^2 + p(1-p)(\lambda_1 - \lambda_0)^2.\end{aligned}$$

[2] See 1

Next,

$$\operatorname{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \mathbb{E}[T] - pm,$$
$$\mathbb{E}[T] = \mathbb{E}[\mathbb{E}[XY \mid X]] = p\,\mathbb{E}[Y \mid X=1] = p\lambda_1,$$

so $\operatorname{Cov}(X,Y) = p(1-p)(\lambda_1 - \lambda_0)$.

Since $T = XY$ takes the value Y when $X=1$ and 0 when $X=0$, we have

$$\mathbb{E}[T^2] = \mathbb{E}[XY^2] = p\,\mathbb{E}[Y^2 \mid X=1] = p\big(\operatorname{Var}(Y \mid X=1) + (\mathbb{E}[Y \mid X=1])^2\big) = 2p\lambda_1^2,$$

hence

$$\operatorname{Var}(T) = \mathbb{E}[T^2] - (\mathbb{E}[T])^2 = 2p\lambda_1^2 - (p\lambda_1)^2 = p(2-p)\lambda_1^2.$$

Finally, $XT = X^2Y = XY = T$, so

$$\operatorname{Cov}(X,T) = \mathbb{E}[XT] - \mathbb{E}[X]\mathbb{E}[T] = \mathbb{E}[T] - p(p\lambda_1) = p(1-p)\lambda_1.$$

Define

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i, \qquad \bar{Y}_n = \frac{1}{n}\sum_{i=1}^n Y_i, \qquad \text{and} \qquad \bar{T}_n = \frac{1}{n}\sum_{i=1}^n X_iY_i.$$

Write a central limit theorem for each of the following quantities in the form

$$\sqrt{n}(Z_n - \mu) \rightsquigarrow \mathcal{N}(0,\sigma^2) \quad \text{or} \quad \sqrt{n}(Z_n - \mu) \rightsquigarrow \mathcal{N}(0,\Sigma),$$

depending on whether Z_n is a random variable or a random vector.

2. $Z_n = \begin{pmatrix} \bar{X}_n \\ \bar{T}_n \end{pmatrix}$.

   Solution. Since $(\bar{X}_n, \bar{T}_n)$ is the sample mean of the i.i.d. vectors (X_i, X_iY_i), the CLT applies. The mean is $\mu = (p, p\lambda_1)$ and the covariance matrix is

   $$\Sigma_{X,XY} = \begin{pmatrix} p(1-p) & p(1-p)\lambda_1 \\ p(1-p)\lambda_1 & p(2-p)\lambda_1^2 \end{pmatrix}.$$

   Therefore,

   $$\sqrt{n}\left(\begin{pmatrix} \bar{X}_n \\ \bar{T}_n \end{pmatrix} - \begin{pmatrix} p \\ p\lambda_1 \end{pmatrix}\right) \rightsquigarrow \mathcal{N}(0, \Sigma_{X,XY}).$$

3. $Z_n = \log(\bar{Y}_n)$.

   Solution. By the CLT, $\sqrt{n}(\bar{Y}_n - m) \rightsquigarrow \mathcal{N}(0, \operatorname{Var}(Y))$.

   Let $g(y) = \log y$. Then $g(m) = \log m$ and $g'(m) = 1/m$. By the delta method,

   $$\sqrt{n}(\log \bar{Y}_n - \log m) \rightsquigarrow \mathcal{N}\left(0, \frac{\operatorname{Var}(Y)}{m^2}\right).$$

Define the average lifetime among passed bulbs

$$\widehat{\lambda}_{1,n} = \begin{cases} \dfrac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n X_i} = \dfrac{\bar{T}_n}{\bar{X}_n}, & \text{if } \sum_{i=1}^n X_i > 0, \\ 0, & \text{if } \sum_{i=1}^n X_i = 0, \end{cases}$$

and the corresponding rate estimator

$$\widehat{\rho}_{1,n} = \begin{cases} \dfrac{1}{\widehat{\lambda}_{1,n}} = \dfrac{\bar{X}_n}{\bar{T}_n}, & \text{if } \sum_{i=1}^n X_i > 0, \\ 0, & \text{if } \sum_{i=1}^n X_i = 0. \end{cases}$$

(Note that $\mathbb{P}(\sum_{i=1}^n X_i = 0) = (1-p)^n \to 0$, so this convention does not affect any CLT/delta-method limits.)
Write a CLT for each of the following choices of Z_n:

4. $Z_n = \widehat{\lambda}_{1,n}$.

   Solution. We use the vector CLT from the previous part and apply the delta method to $g(\bar{X}_n, \bar{T}_n)$, with $\mu = (p, p\lambda_1)$ and covariance $\Sigma_{X,XY}$.

   Let $g(s,t) = t/s$. Then $g(\mu) = \lambda_1$ and $\nabla g(s,t) = \big(-t/s^2,\ 1/s\big)$, so $\nabla g(\mu) = (-\lambda_1/p,\ 1/p)$. Therefore,

   $$\sqrt{n}(\widehat{\lambda}_{1,n} - \lambda_1) \rightsquigarrow \mathcal{N}\left(0, \frac{\lambda_1^2}{p}\right).$$

5. $Z_n = \widehat{\rho}_{1,n}$.

   Solution. Let $g(s,t) = s/t$. Then $g(\mu) = 1/\lambda_1$ and $\nabla g(s,t) = \big(1/t,\ -s/t^2\big)$, so $\nabla g(\mu) = \big(1/(p\lambda_1),\ -1/(p\lambda_1^2)\big)$. Therefore,

   $$\sqrt{n}\left(\widehat{\rho}_{1,n} - \frac{1}{\lambda_1}\right) \rightsquigarrow \mathcal{N}\left(0, \frac{1}{p\lambda_1^2}\right).$$

Extra probability practice (not graded)

Problem 6. Let X be a random variable with pmf given by

$$\mathbb{P}(X = k) = \frac{c\lambda^k}{k!}, k = 0, 1, 2, \ldots$$

for some $\lambda > 0$.

1. What is the value of c? (a) 1 (b) λ (c) $\boldsymbol{e^{-\lambda}}$ (d) e^{λ}
2. What is $\mathbb{E}[X]$? (a) 1 (b) $\boldsymbol{\lambda}$ (c) $e^{-\lambda}$ (d) e^{λ}
3. What is $\mathbb{V}[X]$? (a) 1 (b) $\boldsymbol{\lambda}$ (c) $e^{-\lambda}$ (d) e^{λ}

Let X be a uniform random variable in the interval $[2, 8]$.

4. What is $\mathbb{E}[X]$?

   (a) 2 (b) 3 (c) **5** (d) 8

5. What is $\mathbb{V}[X]$?

   (a) 2 (b) **3** (c) 5 (d) 8

6. What[3] is $\mathbb{P}[\log(X) \leq 1]$ approximately?

   (a) **.12** (b) 0.8 (c) $-.1$ (d) 0

Let X be an exponential random variable with parameter 3 and Y be a Poisson random variable with parameter 2.[4] Assume that X and Y are independent.

7. What is $\mathbb{E}[X^2 + Y^2]$?

   (a) 12 (b) 23 (c) **24** (d) 36

8. What is $\mathbb{E}[X^2 Y]$?

   (a) 12 (b) 23 (c) 24 (d) **36**

9. What is $\mathbb{V}(2X + 3Y)$?

   (a) 24 (b) 34 (c) 44 (d) **54**

Let $X \geq 0$ be a positive random variable such that $\mathbb{E}[X] = \lambda$.

10. Which is correct?

    (a) $\mathbb{E}[1/X] = 1/\lambda$ (b) $\boldsymbol{\mathbb{E}[1/X] \geq 1/\lambda}$ (c) $\mathbb{E}[1/X] \leq 1/\lambda$

Let X and Y be two random variables such that X is a Bernoulli random variable with parameter $p \in (0, 1)$, and $Y^2 + 2XY = 3X^2$ almost surely.

[3] all logs are natural (base e) unless specified otherwise

[4] See 1

11. What is $\mathbb{E}[Y]$?

    (a) 0 (b) $-3p$ (c) X (d) $-3X$ (e) **Some number in $[-3, 1]$**

Let X and Y be two independent, identically distributed random variables.

12. Compute the conditional expectation $\mathbb{E}[X|X+Y=x]$.

    (a) $\boldsymbol{x/2}$ (b) x (c) $-x$ (d) 0