18.650. Fall 2025
Prof. Rigollet

Date: 9/26/2025

Name:________________________________

Test 1 (Solutions)

PLEASE DO NOT TURN THIS PAGE OR START ANSWERING QUESTIONS UNTIL YOU ARE INSTRUCTED TO DO SO.

1. You are allowed a two-sided letter-sized cheat sheet.
2. No other notes or books are allowed
3. Calculators and connected devices like phones, laptops, or tablets are strictly forbidden.
4. The test starts at 10:05
5. The test ends at 10:55 regardless of your time of arrival.
6. All questions should be answered on the present exam sheet
7. Make sure to mark your name on the first page and **do not remove the staple**

Problem1	30	
Problem2	75	
TOTAL	100	

Problem 1. (30 pts) Check the (unique) correct answer. No justification is required. 5 points each.

1. An estimator of θ whose MSE is equal to 0 is...

   ◯ unbiased

   ◯ deterministic

   ◯ equal to θ almost surely

   √ **All of the above**

   Recall that MSE = bias^2 + variance, so MSE being 0 means that both the bias and variance must be 0.

2. In a histogram with *unequal* bin widths, which statement is true?

   ◯ Taller bars always correspond to larger proportions of observations.

   √ **The *area* of each bar equals the proportion of observations in its bin.**

   ◯ The *height* of each bar equals the proportion in its bin regardless of width.

   ◯ The *area* of each bar is equal to $1/k$ if there are k bins.

3. Which pair of summary statistics is most *robust* to outliers?

   ◯ mean and standard deviation

   √ **median and interquartile range (IQR)**

   ◯ mean and IQR

   ◯ median and standard deviation

4. Let X be the time (in minutes) it takes to find a free dock for a Blue Bike on MIT campus. Assume $X \sim \mathcal{N}(10, 4)$. What is the probability that it takes more than 9 minutes to dock a bike?

   - $\surd$ 0.6915
   - ◯ 0.3085
   - ◯ 0.5987
   - ◯ 0.5000

   Note that X has the same distribution as $2Z + 10$, where $Z \sim \mathcal{N}(0, 1)$. So

$$\mathbb{P}(X > 9) = \mathbb{P}(Z > -0.5) = \Phi(0.5) \approx 0.6915.$$

5. Let $X_1, \ldots, X_n \sim \text{Bernoulli}(1/4)$ be i.i.d., and let $\hat{p} = \bar{X}_n$. Define $g(p) = 2\sqrt{p}$. By the Delta Method,

$$\sqrt{n}\left(g(\hat{p}) - g(1/4)\right) \rightsquigarrow \mathcal{N}(0, \sigma^2).$$

   What is σ^2?

   - ◯ .25
   - ◯ 0.5
   - $\surd$ .75
   - ◯ 1

   We have $\mathbb{E}[X_i] = p = \frac{1}{4}$ and $\mathbb{V}(X_i) = p(1-p) = \frac{3}{16}$, so by CLT we have

$$\sqrt{n}(\hat{p} - \frac{1}{4}) \rightsquigarrow \mathcal{N}(0, \frac{3}{16}).$$

   Now, applying Delta method, we have that

$$\sqrt{n}\left(g(\hat{p}) - g(1/4)\right) \rightsquigarrow \mathcal{N}(0, \sigma^2),$$

   where $\sigma^2 = \frac{3}{16} \cdot g'(p)^2$. We have $g'(p) = p^{-1/2}$, so $\frac{3}{16} \cdot g'(p)^2 = \frac{3}{16} \cdot 4 = \frac{3}{4} = .75$.

6. Let $Z = (Z_1, Z_2)^\top$ be bivariate normal with mean $\mu = (1, 1)^\top$ and covariance $\Sigma = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. What is the distribution of $Y = Z_1 - 2Z_2$?

   √ $\mathcal{N}(-1, 6)$

   ◯ $\mathcal{N}(-1, 10)$

   ◯ $\mathcal{N}(0, 6)$

   ◯ $\mathcal{N}(-1, 5)$

   Note that $Y = (1, -2) \cdot (Z_1, Z_2)^\top$, so that

$$\mathbb{E}[Y] = (1, -2) \cdot (1, 1)^\top = -1$$

   and

$$\mathbb{V}(Y) = (1, -2) \cdot \Sigma \cdot (1, -2)^\top = 6.$$

   Since Y must be normal (as it is a linear combination of normals), we have that $Y \sim \mathcal{N}(-1, 6)$.

Problem 2. (70 pts + 5 bonus pts)

Let X be a random vector with unknown mean $\mu = (a, b)^\top \in \mathbb{R}^2$ and covariance matrix

$$\Sigma = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Define $Y^+ = X^{(1)} + X^{(2)}$ and $Y^- = X^{(1)} - X^{(2)}$, where $X = (X^{(1)}, X^{(2)})^\top$.

Our goal in this exercise is to estimate the parameter $\theta = a^2 - b^2$. To that end, we observe $X_1, \ldots, X_n$ which are i.i.d. with the same distribution as X and we propose to use the estimator

$$\hat{\theta} = (\bar{Y}_n^+)(\bar{Y}_n^-),$$

where

$$\bar{Y}_n^\pm = \frac{1}{n}\sum_{i=1}^{n}(X_i^{(1)} \pm X_i^{(2)}).$$

1. (10 points) What is the mean of $(Y^+, Y^-)^\top$?

   Note that $(Y^+, Y^-)^\top = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \cdot X$. Therefore,

   $$\mathbb{E}[(Y^+, Y^-)^\top] = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \mu = \boxed{(a+b, a-b)^\top}.$$

2. (10 points) What is the covariance matrix of $(Y^+, Y^-)^\top$?

   Using the fact that $(Y^+, Y^-)^\top = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \cdot X$, we have that the covariance matrix is

   $$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \Sigma \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^\top = \boxed{\begin{pmatrix} 6 & 0 \\ 0 & 2 \end{pmatrix}}.$$

3. (10 points) Is $\hat{\theta}$ consistent? Why?

   By LLN,

   $$\bar{Y_n}^\pm \xrightarrow{\mathbb{P}} \mathbb{E}[X^{(1)} \pm X^{(2)}] = a \pm b.$$

   Therefore, using the product property of convergence in probability, we have that

   $$(\bar{Y_n}^+)(\bar{Y_n}^-) \xrightarrow{\mathbb{P}} (a+b)(a-b) = a^2 - b^2 = \theta.$$

   Hence, $\hat{\theta} \xrightarrow{\mathbb{P}} \theta$, so $\hat{\theta}$ is consistent.

4. (10 points) Compute the bias of $\hat{\theta}$. Is it biased or unbiased?
Denote $X_i^{(1)} \pm X_i^{(2)}$ as $Y_i^{\pm}$.
Solution 1:

$$\begin{aligned}
\mathbb{E}[\hat{\theta}] &= \mathbb{E}[\frac{1}{n}\sum_{i=1}^{n} Y_i^+ \cdot \frac{1}{n}\sum_{i=1}^{n} Y_i^-] \\
&= \frac{1}{n^2}\mathbb{E}[\sum_{i=1}^{n} Y_i^+ Y_i^- + \sum_{i \neq j} Y_i^+ Y_j^-] \\
&= \frac{1}{n^2}\mathbb{E}[nY_1^+ Y_1^- + n(n-1)Y_1^+ Y_2^-] \\
&= \frac{1}{n^2}[n(a^2-b^2) + n(n-1)(a^2-b^2)] \\
&= a^2 - b^2
\end{aligned}$$

Therefore, $\mathbb{E}[\hat{\theta}] - \theta = 0$ and $\hat{\theta}$ is unbiased.
Solution 2:

$$\begin{aligned}
\mathbb{E}[\hat{\theta}] &= \mathbb{E}[\overline{Y_n^+} \cdot \overline{Y_n^-}] \\
&= \mathbb{E}[\overline{Y_n^+}]\mathbb{E}[\overline{Y_n^-}] + \mathrm{Cov}(\overline{Y_n^+}, \overline{Y_n^-}) \\
&= (a+b)(a-b) \\
&= a^2 - b^2
\end{aligned}$$

where we used the fact that $\mathrm{Cov}(\overline{Y_n^+}, \overline{Y_n^-})$ is 0, because $\mathrm{Cov}(Y_i^+, Y_i^-) = 0$ (from part 2) and $\mathrm{Cov}(Y_i^+, Y_j^-) = 0$ for $i \neq j$ (as they are functions of X_i and X_j, respectively, which are independent). Therefore, $\mathbb{E}[\hat{\theta}] - \theta = 0$ and $\hat{\theta}$ is unbiased.
Solution 3:

$$\begin{aligned}
\mathbb{E}[\hat{\theta}] &= \mathbb{E}[\overline{Y_n^+} \cdot \overline{Y_n^-}] \\
&= \mathbb{E}[(\bar{X}_n^{(1)})^2 - (\bar{X}_n^{(2)})^2] \\
&= \mathbb{V}[\bar{X}_n^{(1)}] + \mathbb{E}[\bar{X}_n^{(1)}]^2 - \mathbb{V}[\bar{X}_n^{(2)}] - \mathbb{E}[\bar{X}_n^{(2)}]^2 \\
&= \frac{2}{n} + a^2 - \frac{2}{n} - b^2 \\
&= a^2 - b^2.
\end{aligned}$$

Therefore, $\mathbb{E}[\hat{\theta}] - \theta = 0$ and $\hat{\theta}$ is unbiased.

5. (10 points) Write a multivariate central limit theorem for $(\bar{Y}_n^+, \bar{Y}_n^-)^\top$.
We can use the multivariate central limit theorem, noting that

$$(\bar{Y_n}^+, \bar{Y_n}^-)^\top = \frac{1}{n}\sum_{i=1}^{n}(Y_i^+, Y_i^-)^\top.$$

So our μ and Σ will be as computed in parts 1 and 2. We have

$$\sqrt{n}((\bar{Y_n}^+, \bar{Y_n}^-)^\top - (a+b, a-b)^\top) \rightsquigarrow \mathcal{N}\left(0, \begin{pmatrix} 6 & 0 \\ 0 & 2 \end{pmatrix}\right).$$

6. (10 points) Show that $\hat{\theta}$ is asymptotically normal and compute its asymptotic variance σ^2.
Note that $\hat{\theta} = g((\bar{Y_n}^+, \bar{Y_n}^-)^\top)$ where $g(x,y) = xy$. We can use the multivariate Delta method, noting the gradient is $\nabla g(x,y) = (y,x)^\top$. We get

$$\sqrt{n}(g(\bar{Y_n}^+, \bar{Y_n}^-)^\top - g(a+b, a-b)^\top) \rightsquigarrow \mathcal{N}\left(0, (a-b, a+b)\begin{pmatrix} 6 & 0 \\ 0 & 2 \end{pmatrix}(a-b, a+b)^\top\right).$$

Simplifying gives

$$\sqrt{n}(\hat{\theta} - \theta) \rightsquigarrow \mathcal{N}\left(0, 8a^2 - 8ab + 8b^2\right).$$

So the asymptotic variance is $8a^2 - 8ab + 8b^2$.

7. (5 points) Propose a consistent estimator $\hat{\sigma}^2$ of σ^2.
Note that $8a^2 - 8ab + 8b^2 = 6(a+b)^2 + 2(a-b)^2$. So $6(\bar{Y_n}^+)^2 + 2(\bar{Y_n}^-)^2$ is a consistent estimator, using the continuous mapping theorem and the addition rule for convergence in probability. Other consistent estimators are possible, such as those that use $\bar{X}_n^{(1)}$ and $\bar{X}_n^{(2)}$.

8. (10 points) Write a 99% confidence interval for θ that is symmetric about $\hat{\theta}$.
Note that $\Phi(2.58) \approx .995$, therefore a 99% confidence interval symmetric about $\hat{\theta}$ would be $\left(\hat{\theta} - 2.58\frac{\hat{\sigma}}{\sqrt{n}}, \hat{\theta} + 2.58\frac{\hat{\sigma}}{\sqrt{n}}\right)$, or

$$\left(\hat{\theta} - 2.58\sqrt{\frac{6(\bar{Y_n}^+)^2 + 2(\bar{Y_n}^-)^2}{n}}, \hat{\theta} + 2.58\sqrt{\frac{6(\bar{Y_n}^+)^2 + 2(\bar{Y_n}^-)^2}{n}}\right).$$

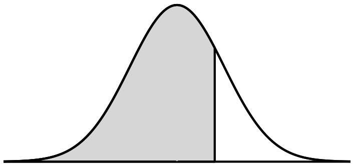

Table 1: The table lists $P(Z \leq z)$ where $Z \sim N(0,1)$ for positive values of z.

	Second decimal place of Z									
Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

*For $Z \geq 3.50$, the probability is greater than or equal to 0.9998.