

IE 241 Engineering Statistics I – Midterm Exam

1, 2, 5, 6 : Graded by Taemin Park

April 18, 2024

3, 4, 7, 8 : Graded by Hyunwoo Cha Time allowed: 2 hr 45 minutes

Name: TA

Student ID: _____

Instructions: Show all your work on your paper; You will NOT receive credits if you do not justify your answers.

1. (10pts) Suppose that there is a 30% chance of rain tomorrow. If it rains, there is a 10% chance classes will be cancelled. If it does not rain, there is only a 2% chance classes will be cancelled.

- (a) (3pts) What is the probability it will rain and classes will be cancelled?

$$\begin{aligned} P(\text{Rain} \& \text{Cancel}) &= P(\text{rain}) \cdot P(\text{cancel} | \text{rain}) = \frac{3}{10} \times \frac{1}{10} = \frac{3}{100} \\ &\quad \text{(+1)} \qquad \qquad \qquad \text{(+1)} \qquad \qquad \qquad \text{(+1)} \end{aligned}$$

- (b) (3pts) What is the total probability (whether or not it rains) that classes will be cancelled?

$$\begin{aligned} P(\text{cancel}) &= P(\text{cancel} \cap \text{rain}) + P(\text{cancel} \cap \text{no rain}) \quad \text{(+1)} \\ &= P(\text{rain}) P(\text{cancel} | \text{rain}) + P(\text{no rain}) P(\text{cancel} | \text{no rain}) \quad \text{(+1)} \\ &= \frac{30}{100} \times \frac{10}{100} + \frac{70}{100} \times \frac{2}{100} = \frac{44}{1000} \quad \text{(+1)} \end{aligned}$$

- (c) (4pts) If classes are indeed cancelled tomorrow, what is the probability that it is due to rain?

$$\begin{aligned} P(\text{rain} | \text{cancel}) &= \frac{P(\text{rain} \cap \text{cancel})}{P(\text{cancel})} = \frac{\frac{3}{100}}{\frac{44}{1000}} = \frac{15}{22} \approx 0.6818 \\ &\quad \text{(+1)} \qquad \qquad \qquad \text{(+1)} \qquad \qquad \qquad \text{(+2)} \end{aligned}$$

2. (10pts) Five balls, numbered 1, 2, 3, 4, and 5, are placed in an urn. Two balls are randomly selected from the five without replacement. Let X be the smaller number of the two.

(a) (5pts) List the possible values for X and assign a probability for each value.

X	1	2	3	4	\Rightarrow Listing (+2)
P_X	$\frac{4}{10}$	$\frac{3}{10}$	$\frac{2}{10}$	$\frac{1}{10}$	\Rightarrow value (+3)

\Rightarrow no partial for each part

(b) (5pts) Let another random variable Y be 1 if $X = 1$, and be 2 if $X > 1$. Find the correlation between X and Y .

$$E(X) = 2, \quad V(X) = 1$$

$$E(Y) = 1.6, \quad V(Y) = 0.24, \quad E(XY) = \frac{2}{5} + \frac{12}{10} + \frac{12}{10} + \frac{8}{10} = 3.6$$

$$\text{Cov}(X, Y) = 3.6 - 3.2 = 0.4 \quad (+2) \text{ with full process}$$

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)} \sqrt{\text{Var}(Y)}} = \frac{0.4}{\sqrt{0.24}} \quad (+1) \quad \text{calculation mistake } (-2)$$

deduction factors are written on exam paper.

3. (6pts) The time (in hours) a manager takes to interview a job applicant has an exponential distribution with $\beta = 1/2$. The applicants are scheduled at quarter-hour intervals, beginning at 8:00 A.M., and the applicants arrive exactly on time.

(a) (3pts) When the applicant with an 8:15 A.M. appointment arrives at the manager's office, what is the probability that she will have to wait before seeing the manager? (Find a numerical answer)

$$\beta = \frac{1}{2}, \quad f(y; \frac{1}{2}) = 2e^{-2y} \quad P(\text{she will have to wait}) = P(\text{First applicant's interview} > \frac{1}{4} \text{ hr})$$

$$\therefore P(Y > \frac{1}{4}) = 1 - P(Y < \frac{1}{4}) = 1 - \int_0^{\frac{1}{4}} 2e^{-2y} dy = 1 - [-e^{-2y}]_0^{\frac{1}{4}} = 1 - (-e^{-\frac{1}{2}} + 1) = e^{-\frac{1}{2}}$$

(b) (3pts) When the applicant with an 8:45 A.M. appointment arrives at the manager's office, how would you find the probability that she will have to wait before seeing the manager? (No need for a numerical answer)

\rightarrow fourth applicant
 $P(\text{The applicant with an 8:45 AM will have to wait})$

$$= P(\text{The total interview time for applicants 1, 2, 3} > \frac{3}{4} \text{ hr})$$

Let $Y^* = Y_1 + Y_2 + Y_3$, where Y_i : interview time for applicant i

$$\text{Then } Y^* \sim \text{Gamma}(3, \frac{1}{2}) \Rightarrow P(Y^* > \frac{3}{4}) = \int_{\frac{3}{4}}^{\infty} f(y^*) dy^* \approx 0.6693$$

(No need for a numerical answer)

4. (15pts) Let X denote the number of digital cameras sold during a particular week by a certain store. The pmf of X is

x	0	1	2	3	4
$p_X(x)$.1	.2	.3	.3	.1

Sixty percent of all customers who purchase these cameras also buy an extended warranty. Let Y denote the number of purchasers during this week who buy an extended warranty.

- (a) (5pts) Find the probability $P(X = 4, Y = 2)$.

$$Y|X \sim B(X, 0.6), \quad P(X, Y) = P(Y|X)P(X).$$

$$\text{By using it, } P(X=4, Y=2) = P(Y=2|X=4) \cdot P(X=4) \quad (C+3)$$

$$P(Y=2|X=4) = \binom{4}{2} \cdot (0.6)^2 (0.4)^2, \quad P(X=4) = 0.1$$

$$\therefore P(X=4, Y=2) = \binom{4}{2} (0.6)^2 (0.4)^2 (0.1) = 0.03456 \quad (\text{no need for calculation})$$

(C+2)

- (b) (5pts) Find the probability $P(Y = 2)$.

$$P(Y=2) = \sum_{i=0}^4 P(X=i, Y=2) = \sum_{i=2}^4 P(X=i, Y=2) \quad (C+3)$$

$$\begin{aligned} &= P(X=2, Y=2) + P(X=3, Y=2) + P(X=4, Y=2) \\ \text{by (a)} \quad &= \binom{2}{2} (0.6)^2 (0.4)^0 (0.3) + \binom{3}{2} (0.6)^2 (0.4)^1 (0.3) + \binom{4}{2} (0.6)^2 (0.4)^2 (0.1) \end{aligned}$$

(C+2)

$$= 0.27216 \quad (\text{no need for calculation})$$

- (c) (5pts) Are X and Y independent? Give a reason for your answer.

Not independent. (C+2)

$$P(X=4)P(Y=2) = (0.1) \cdot (0.27216) \neq P(X=4, Y=2)$$

or

$$P(X=0, Y=2) = 0 \neq P(X=0)P(Y=2).$$

\Rightarrow you only need to see one of many cases like this.

(C+3)

5. (15pts) Suppose Y_1 and Y_2 have a bivariate normal distributed with $N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. Let $\sigma_{12} = \rho\sigma_1\sigma_2$.

(a) (5pts) Show that $X := Y_1 - aY_2$ is uncorrelated with Y_2 if $a = \sigma_{12}/\sigma_2^2$. Explain why this implies that X and Y_2 are independent.

$$\begin{aligned} \text{cov}(X, Y_2) &= \text{cov}(Y_1, Y_2) - a \text{var}(Y_2) \\ &= \sigma_{12} - a\sigma_2^2, \text{ when } a = \frac{\sigma_{12}}{\sigma_2^2} \end{aligned}$$

$$\text{cov} = 0 \Leftrightarrow \text{corr} = 0 \quad (C+3)$$

(Y_1, Y_2) are jointly normal $\Rightarrow (X, Y_2)$ are jointly normal $(C+2)$

(b) (5pts) Derive $E(Y_1|Y_2)$ from (a). (Hint: Use $E(Y_1|Y_2) = E(X|Y_2) + aE(Y_2|Y_2)$).

$$E(Y_1|Y_2) = E(X|Y_2) + aE(Y_2|Y_2)$$

$$\begin{aligned} &\downarrow (C+2) \quad \downarrow (C+2) \\ &= E(X) + aY_2 \end{aligned}$$

$$= \mu_1 - a\mu_2 + aY_2 = \mu_1 + a(Y_2 - \mu_2) \quad (C+1)$$

(c) (5pts) Derive $\text{Var}(Y_1|Y_2)$ from (a).

$$\text{Var}(Y_1|Y_2) = \text{Var}(X + aY_2 | Y_2)$$

$$= \text{Var}(X|Y_2) + a^2 \text{Var}(Y_2|Y_2) \quad \text{since } \text{cov}(X, Y_2) = 0 \quad (C+2)$$

$$= \text{Var}(X) \quad (C+1)$$

$$= \text{Var}(Y_1) + a^2 \text{Var}(Y_2) - 2a \text{cov}(Y_1, Y_2) \quad (C+1)$$

$$= \sigma_1^2 + a^2 \sigma_2^2 - 2a \sigma_{12}$$

$$= \sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2} \quad (C+1)$$

6. (16pts) Suppose that Y_1, \dots, Y_n are i.i.d. random variables from $\text{Poisson}(\lambda)$.

(a) (3pts) Find the distribution of $T = Y_1 + \dots + Y_n$ by finding mgf of T .

$$\text{mgf of } T: E(e^{tT}) \stackrel{(+1)}{=} \prod_{i=1}^n E(e^{tY_i}) \stackrel{(+1)}{=} \exp(n\lambda(e^t - 1))$$

$$\Rightarrow \text{mgf of } \text{poisson}(n\lambda) \quad (+1)$$

(b) (3pts) Let $\bar{Y} = T/n$. show that the conditional distribution of Y_1 given that $\bar{Y} = m$ is a binomial distribution. Specify the parameters.

$$\begin{aligned} P(Y_1 = y \mid \bar{Y} = m) &= P(Y_1 = y \mid T = nm) = \frac{P(Y_1 = y) P(Y_2 + \dots + Y_n = nm - y)}{P(T = nm)} \quad (+1) \\ &= \dots = \binom{nm}{y} \cdot \left(\frac{\lambda}{n\lambda}\right)^y \cdot \left(\frac{(n-1)\lambda}{n\lambda}\right)^{nm-y} \sim B(nm, \frac{1}{n}) \quad (+1) \end{aligned}$$

(c) (3pts) Using (b), find the conditional mean $E(Y_1 \mid \bar{Y})$ and conditional variance $\text{Var}(Y_1 \mid \bar{Y})$.

$$E(Y_1 \mid \bar{Y}) = nm \times \frac{1}{n} = m = \bar{Y} \quad (+1)$$

$$V(Y_1 \mid \bar{Y}) = nm \times \frac{1}{n} \times \frac{n-1}{n} = m(1 - \frac{1}{n}) = \bar{Y} \cdot (1 - \frac{1}{n}) \quad (+2)$$

(d) (4pts) Using (c), show that $E(S^2 \mid \bar{Y}) = \bar{Y}$.

$$\begin{aligned} E(S^2 \mid \bar{Y}) &= \frac{1}{n-1} E(\sum_{i=1}^n Y_i^2 - n\bar{Y}^2 \mid \bar{Y}) \quad (+1) \\ &= \frac{1}{n-1} (n E(Y_1^2 \mid \bar{Y}) - n\bar{Y}^2) \quad (+1) \\ &= \frac{1}{n-1} (n [\text{Var}(Y_1 \mid \bar{Y}) + (E(Y_1 \mid \bar{Y}))^2] - n\bar{Y}^2) \quad (+1) \\ &= \frac{1}{n-1} (n \bar{Y} (1 - \frac{1}{n}) + n\bar{Y}^2 - n\bar{Y}^2) = \bar{Y} \quad (+1) \end{aligned}$$

(e) (3pts) Using (d), show that $\text{Var}(S^2) \geq \text{Var}(\bar{Y})$.

$$\text{Var}(S^2) = E(\text{Var}(S^2 \mid \bar{Y})) + \text{Var}(E(S^2 \mid \bar{Y})) \quad (+1)$$

$$\geq \text{Var}(E(S^2 \mid \bar{Y})) \stackrel{(+1)}{=} \text{Var}(\bar{Y})$$

$$\text{since } \text{Var}(S^2 \mid \bar{Y}) \geq 0 \Rightarrow E(\text{Var}(S^2 \mid \bar{Y})) \geq 0 \quad (+1)$$

7. (16pts) Let X and Y have a joint density function given by

$$f(x, y) = 3x, \quad 0 \leq y \leq x \leq 1.$$

(a) (4pts) Find $P(X < a)$ for $0 < a < 1$.

$$f(x) = \int_0^x 3x \, dy = [3xy]_0^x = 3x^2$$

$$\therefore P(X < a) = \int_0^a 3x^2 \, dx = a^3 \quad (0 < a < 1) \quad (C+4)$$

(b) (4pts) Show that X has a beta distribution. Identify the parameters.

$$\text{pdf of beta}(\alpha, \beta) : f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot x^{\alpha-1} \cdot (1-x)^{\beta-1}$$

$$\text{by (a), } f(x) = 3x^2 = \frac{\Gamma(4)}{\Gamma(3)\Gamma(1)} \cdot x^{3-1} \cdot (1-x)^{1-1} = 3x^2 \quad (C+3)$$

$$\therefore X \sim \text{beta}(3, 1) \quad (C+1)$$

(c) (4pts) Show that given $X = x$, Y has a uniform distribution **and** find the conditional mean of Y given $X = x$.

$$f_{Y|X}(x, y) = \frac{f_{X,Y}(x, y)}{f_X(x)} = \frac{3x}{3x^2} = \frac{1}{x} \quad (0 \leq y \leq x)$$

$$\sim \text{Uniform}(0, x) \quad (C+2)$$

$$\therefore E(Y|X=x) = \frac{x}{2} \quad (\text{using mean of uniform distribution}) \quad (C+2)$$

(d) (4pts) Using (b) and (c), find the mean of Y .

$$E(Y) = E(E(Y|X)) = E\left(\frac{X}{2}\right) = \frac{1}{2} E(X) = \frac{3}{8} \quad (C+4)$$

by (c)

by (b), mean of $\text{beta}(3, 1) = \frac{3}{3+1}$

8. (12pts) Suppose a random variable X has a probability density function

$$f(x) = ke^{-2(x-1)^2}, \quad -\infty < x < \infty.$$

(a) (3pts) Find k .

Let $X \sim N(1, (\frac{1}{2})^2)$, pdf of X : $f(x) = \left(\frac{\pi}{2}\right)^{-\frac{1}{2}} \cdot \exp^{-2(x-1)^2}$

$$\therefore k = \left(\frac{\pi}{2}\right)^{-\frac{1}{2}} \quad (+3)$$

(b) (3pts) Find the probability $P(X < 2)$.

$$P(X < 2) = P\left(\frac{X-1}{\frac{1}{2}} < \frac{2-1}{\frac{1}{2}}\right) \quad (+2)$$

$$= P(Z < 2) \approx 0.9772 \quad (\text{by } Z\text{-table}) \quad (+2)$$

(c) (3pts) Find the 80% percentile of the distribution of X .

$P(X < c) = 0.8$. we need to find c

$$\Leftrightarrow P\left(\frac{X-1}{\frac{1}{2}} < \frac{c-1}{\frac{1}{2}}\right) = P(Z < 2(c-1)) = 0.8 \quad (+2)$$

by Z -table $2(c-1) = 0.84$ or $0.85 \Rightarrow c = 1.42$ or 1.425 (+2)

(d) (3pts) Find $E((X-1)^4)$.

sol 1) $\left(\frac{X-1}{\frac{1}{2}}\right)^2 \sim \chi^2(1) \Rightarrow E(4(X-1)^2) = 1, \quad V(4(X-1)^2) = 2$

$$\Rightarrow E((X-1)^2) = \frac{1}{4}, \quad V((X-1)^2) = \frac{1}{8}$$

$$\therefore E((X-1)^4) = V((X-1)^2) + [E((X-1)^2)]^2 = \frac{3}{16}$$

sol 2) Let $Y = X-1, Y \sim N(0, (\frac{1}{2})^2)$, we need to calculate $E(Y^4)$

using mgf $m_Y(t)$. $E(Y^4) = m_Y^{(4)}(0) = \frac{3}{16}$

sol 3) Directly calculate $\int_{-\infty}^{\infty} (x-1)^4 f(x) dx$. sol 1 or 2 or 3 (+2)

\Rightarrow No matter how you calculate it, the answer should be correct. (+1)

Distribution	PDF or PMF	Mean	Variance	Moment-Generating Function
Continuous distributions				
Uniform	$f(y) = \frac{1}{\theta_2 - \theta_1}$	$\frac{\theta_1 + \theta_2}{2}$	$\frac{(\theta_2 - \theta_1)^2}{12}$	$\frac{e^{t\theta_2} - e^{t\theta_1}}{t(\theta_2 - \theta_1)}$
Normal	$f(y) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left[-\left(\frac{1}{2\sigma^2}\right)(y - \mu)^2\right]$	μ	σ^2	$\exp\left(\mu t + \frac{t^2\sigma^2}{2}\right)$
Exponential	$f(y) = \frac{1}{\beta} e^{-y/\beta}$	β	β^2	$(1 - \beta t)^{-1}$
Gamma	$f(y) = \frac{1}{\Gamma(\alpha)\beta^\alpha} y^{\alpha-1} e^{-y/\beta}$	$\alpha\beta$	$\alpha\beta^2$	$(1 - \beta t)^{-\alpha}$
Chi-square	$f(y) = \frac{(y)^{(\nu/2)-1} e^{-y/2}}{2^{\nu/2}\Gamma(\nu/2)}$	ν	2ν	$(1 - 2t)^{-\nu/2}$
Beta	$f(y) = \left[\frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)}\right] y^{\alpha-1} (1 - y)^{\beta-1}$	$\frac{\alpha}{\alpha + \beta}$	$\frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$	does not exist in closed form
Bivariate Normal	$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}Q\right),$ $Q = \left(\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}\right)$			
Discrete distributions				
Binomial	$p(y) = \binom{n}{y} p^y (1-p)^{n-y}$	np	$np(1-p)$	$[pe^t + (1-p)]^n$
Geometric	$p(y) = p(1-p)^{y-1}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1 - (1-p)e^t}$
Hypergeometric	$p(y) = \frac{\binom{r}{y} \binom{N-r}{n-y}}{\binom{N}{n}}$	$\frac{nr}{N}$	$n\left(\frac{r}{N}\right)\left(\frac{N-r}{N}\right)\left(\frac{N-n}{N-1}\right)$	
Poisson	$p(y) = \frac{\lambda^y e^{-\lambda}}{y!}$	λ	λ	$\exp[\lambda(e^t - 1)]$
Negative binomial	$p(y) = \binom{y-1}{r-1} p^r (1-p)^{y-r}$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$	$\left[\frac{pe^t}{1 - (1-p)e^t}\right]^r$
Multinomial	$p(y_1, \dots, y_k) = \frac{n!}{y_1!y_2!\dots y_k!} p_1^{y_1} \dots p_k^{y_k}$			

Standard Normal Probabilities

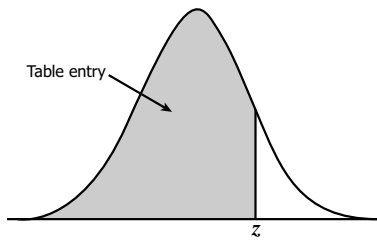


Table entry for z is the area under the standard normal curve to the left of z .

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998