IE 241 Engineering Statistics I – Final Exam

· Claim

June 13, 2024 $\quad$ 1, 3, 5 → Taemin park $\quad$ Time allowed: 2 hr 45 minutes

2, 4, 6 → Hyunwoo Cha

Name: TA $\quad$ **Student ID:** ______________________

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

1. (25pts) A salesperson wants to estimate the probability of a sale on a single contact. From her previous record, the probability would not exceed 20%.

(a) (5pts) Let Y denote the number of sales from 50 independent contacts. Find a 95% confidence interval for the probability of a sale on a single contact. Show that your interval has at least 95% coverage of the true parameter.

$$\hat{p} = \frac{Y}{50}, \quad \hat{p} \pm 1.96 \cdot \sqrt{\frac{p(1-p)}{p}} \simeq \hat{p} \pm 1.96 \sqrt{\frac{\hat{p}(1-\hat{p})}{50}}$$

$$\hat{p} \le 0.2 \Rightarrow \sqrt{\frac{\hat{p}(1-\hat{p})}{50}} \le \sqrt{\frac{0.2 \cdot 0.8}{50}}$$

$\Rightarrow$ C.I is wider with $\sqrt{\frac{0.2 \cdot 0.8}{50}} \Rightarrow \hat{p} \pm 1.96 \sqrt{\frac{0.2 \times 0.8}{50}}$

has at least 95% coverage of true parameter

(b) (5pts) Find a desired sample size if she wants the width of the confidence interval to be no more than 4%.

$$2 \cdot 1.96 \sqrt{\frac{0.2 \cdot 0.8}{n}} \le 0.04 \Rightarrow n \ge \left(\frac{1.96}{0.02}\right)^2 \cdot 0.2 \cdot 0.8$$

$$= 1536.64$$

mistake (-1)

(c) (5pts) Let θ be the probability that she makes sales from **two** consecutive contacts. Find an unbiased estimator of θ.

$\theta = p^2$, $E(Y^2) = V(Y) + E(Y)^2 = np(1-p) + n^2p^2$ $\left(p = \frac{Y}{n}\right)$

(+1)

$$\therefore \hat{\theta} = \frac{Y(Y-1)}{n(n-1)}$$ (+4)

(d) (5pts) Find a MLE of θ.

$$L(y_1, \cdots, y_n \mid p) = p^{\Sigma y_i} \cdot (1-p)^{n-\Sigma y_i}$$ (+1)

$\hat{p}^{MLE} = \frac{Y}{n}$, $\hat{\theta}^{MLE} = (\hat{p}^{MLE})^2 = \left(\frac{Y}{n}\right)^2$ (+2)

(+2)

(e) (5pts) Find an approximate 95% confidence interval for θ for large n.

$I(p) = E\left(-\frac{d^2}{dp^2}(\ln f(y|p))\right) = \frac{1}{p(1-p)}$ or $V(\hat{\theta}^{MLE}) = 2p^2 \cdot \frac{p(1-p)}{n}$

(+1) or (+1)

$(\hat{p}^{MLE})^2 \sim N\left(p^2, \frac{4p^2}{nI(p)}\right)$ (+2)

$\Rightarrow \left(\frac{Y}{n}\right)^2 \pm 1.96 \cdot \sqrt{\frac{4\hat{p}^3(1-\hat{p})}{n}}$ $\left(\hat{p} = \frac{Y}{n}\right)$

(+2)

* Most students solved the problem by confusing the liklihood and probability.
But two concepts are different.

2. (10pts) An urn has five balls inside. Each ball is either red or blue. It is unknown how many balls of each color are in the urn.

(a) (5pts) Three balls are randomly drawn **with** replacement, and two red balls and one blue ball are selected. Find the MLE of the number of red balls.

$L(\theta \mid 2\text{Red}, 1\text{Blue}) = \binom{3}{2}\left(\frac{\theta}{5}\right)^2\left(1-\frac{\theta}{5}\right)^1$ is maximized

when $\theta = 3.xx$

$L(3) > L(4) \Rightarrow \hat{\theta}^{MLE} = 3$

< majority of deduction factors >

- invalid liklihood equation (−2)
- invalid liklihood calculation (−1)
- MLE is not natural number (−1)

(b) (5pts) Three balls are randomly drawn **without** replacement, and two red balls and one blue ball are selected. Find the MLE of the number of red balls.

$$L(\theta \mid 2\text{red}, 1\text{blue}) = \frac{\binom{\theta}{2}\binom{5-\theta}{1}}{\binom{5}{3}}$$

$L(3) = L(4) = \frac{6}{10}$, $L(2) = \frac{3}{10}$

$\therefore \hat{\theta}^{MLE} = 3$ or 4

< majority of deduction factors >

- invalid liklihood equation (−2)
- invalid liklihood calculation (−1)
- MLE is not natural number (−1)

3. (5pts) Suppose that X_1, X_2 are i.i.d. from the uniform distribution $U(0, \theta)$. Find the distribution of $Y = \sqrt{X_1 X_2}$.

$$Y = \sqrt{X_1 X_2}, \quad Z = X_1 \Rightarrow \left[\begin{array}{l} X_1 = Z \\ X_2 = \frac{Y^2}{Z} \end{array} \right., \quad |J| = \begin{vmatrix} 0 & 1 \\ \frac{2y}{z} & -\frac{y^2}{z^2} \end{vmatrix} = \frac{2y}{z}$$

(+1)

$$f_{Y,Z}(y,z) = \frac{1}{\theta^2} \cdot \frac{2y}{z}, \quad 0 < \frac{y^2}{z} < \theta, \quad 0 < z < \theta \Rightarrow 0 < y^2 < z\theta < \theta^2$$

$$\Rightarrow \frac{y^2}{\theta} < z < \theta, \quad 0 < y < \theta$$ (+2)

$$f_Y(y) = \int_{\frac{y^2}{\theta}}^{\theta} \frac{1}{\theta^2} \cdot \frac{2y}{z} \, dz = \left[\frac{2y}{\theta^2} \cdot \ln z \right]_{\frac{y^2}{\theta}}^{\theta} = \frac{2y}{\theta^2} \cdot \ln \left(\frac{\theta^2}{y^2} \right) \quad (0 < y < \theta)$$

(+2)

4. (5pts) Suppose you have a random number generator that can generate an i.i.d. sample from a uniform distribution $U(0, 1)$. Describe, in as much detail as possible, how you can use this generator to produce an i.i.d. sample from a $Beta(3, 1)$.

Let $Y = X^{\frac{1}{\alpha}} \quad (0 \le X \le 1)$ (because $0 \le X \le 1$)

$$F_Y(y) = P(Y \le y) = P(X^{\frac{1}{\alpha}} \le y) = P(X \le y^\alpha)$$
$$= y^\alpha \quad (0 \le y \le 1)$$

$f_Y(y) = \alpha y^{\alpha-1}$. pdf of $Beta(\alpha, 1) = \alpha y^{\alpha-1}$

$\therefore$ generate X from $U(0,1)$ then $Y = X^{\frac{1}{3}}$, $Y \sim Beta(3,1)$

(+5)

5. (20pts) Suppose that X_1 and X_2 be i.i.d. random variables from standard normal distribution with pdf $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$, and cdf $\Phi(x) = \int_{-\infty}^{x} f(y)dy$.

(a) (5pts) Find the moment generating function of $Y = X_1^2$, that is, $\mathrm{E}(e^{tX_1^2})$, to show that Y has a chi-square distribution with degrees of freedom 1. Make sure you show your work.

$$E(e^{tx^2}) = \int \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(1-2t)x^2} dx = (1-2t)^{-\frac{1}{2}} \cdot \underbrace{\int \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{k^2}{2}} \cdot dk}_{=1} = (1-2t)^{-\frac{1}{2}}$$

$\sqrt{1-2t}\, x = k$ (+3)

$\Rightarrow$ mgf of $\chi^2(1)$ (+2)

(b) (10pts) Let $Y_1 = \max(X_1, X_2)$. Show that Y_1^2 also has a chi-square distribution with degrees of freedom 1. (Hint: Find the cdf of Y_1^2. You may need to use $\Phi(x) = 1 - \Phi(-x)$, and $\phi(x) = \phi(-x)$.)

$$P(Y^2 \le y) = P(-\sqrt{y} \le Y \le \sqrt{y}) = P(Y \le \sqrt{y}) - P(Y \le -\sqrt{y})$$

$$= P(X_1 \le \sqrt{y})^2 - P(X_1 \le -\sqrt{y})^2$$

$$= \Phi(\sqrt{y})^2 - \Phi(-\sqrt{y})^2 \quad (+5)$$

pdf: $2\Phi(\sqrt{y}) \cdot \phi(\sqrt{y}) \cdot \frac{1}{2\sqrt{y}} - 2\Phi(-\sqrt{y}) \cdot \phi(-\sqrt{y}) \cdot \left(-\frac{1}{2\sqrt{y}}\right)$

$$= \Phi(\sqrt{y}) \cdot \phi(\sqrt{y}) \cdot \frac{1}{\sqrt{y}} + (1 - \Phi(\sqrt{y}))\, \phi(\sqrt{y}) \cdot \frac{1}{\sqrt{y}}$$

$$= \phi(\sqrt{y}) \cdot \frac{1}{\sqrt{y}} = \frac{1}{\sqrt{2\pi y}} e^{-\frac{1}{2}y} \quad \sim \chi^2(1)$$

(c) (5pts) Let $Y_2 = \min(X_1, X_2)$. Explain why Y_2^2 also has a chi-square distribution with degrees of freedom 1.

$$[\min(X_1, X_2)]^2 \overset{d}{=} [\min(-X_1, -X_2)]^2 = [-\max(X_1, X_2)]^2 = [\max(X_1, X_2)]^2$$

$$\underbrace{X_1 \overset{d}{=} -X_1, \quad X_2 \overset{d}{=} -X_2}$$

(+5)

6. (35pts) Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal distribution with mean 1 and variance θ. Consider the following four estimators of θ.

$$\hat{\theta}_1 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2, \quad \hat{\theta}_2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2, \quad \hat{\theta}_3 = \frac{1}{n}\sum_{i=1}^{n}(X_i - 1)^2. \quad \hat{\theta}_4 = \frac{1}{n}\sum_{i=1}^{n}X_i^2 - 1.$$

(a) (10pts) Which are unbiased estimators? Justify your answer.

<Grading Criteria>

- $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4 \Rightarrow$ 2.5pts each.
- Choosing the wrong U.E (-1)
- partial credit is awarded for calculation mistake

$$E(\hat{\theta}_1) = \frac{n-1}{n}\theta$$

$$E(\hat{\theta}_2) = E(\hat{\theta}_3) = E(\hat{\theta}_4) = \theta$$

If you want to get full credit, show your work!

(b) (10pts) Find the MSE (Mean Square Error) of these estimators. Which estimator is the best, based on MSE?

< Grading Criteria >

- MSE of $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4 \Rightarrow$ 2 / 2 / 2 / 3 pts each
- Choosing best based on MSE $\Rightarrow$ 1 pts
- partial credit is awarded for calculation mistake

$$V(\hat{\theta}_1) = \left(\frac{n-1}{n}\right)^2 \cdot V(\hat{\theta}_2) = \left(\frac{n-1}{n}\right)^2 \cdot \frac{2\theta^2}{n-1} = \frac{2(n-1)}{n^2} \cdot \theta^2$$

using variance of χ^2-dist

- $$MSE(\hat{\theta}_1) = \frac{2(n-1)}{n^2} \cdot \theta^2 + \frac{\theta^2}{n^2} = \frac{2n-1}{n^2} \cdot \theta^2 \quad (+2)$$
- $$MSE(\hat{\theta}_2) = V(\hat{\theta}_2) = \frac{2}{n-1}\theta^2 \quad (+2)$$

$$\frac{X_i - 1}{\sqrt{\theta}} \sim Z \Rightarrow \frac{\Sigma(X_i-1)^2}{\theta} \sim \chi^2(n)$$

$$V(\Sigma(X_i-1)^2) = 2n\theta^2$$

- $$MSE(\hat{\theta}_3) = V(\hat{\theta}_3) = \frac{1}{n^2} \cdot 2n\theta^2 = \frac{2\theta}{n} \quad (+2)$$

mgf of X_1: $\exp\left(t + \frac{1}{2}\theta t^2\right)$

$$\rightarrow m'(t) = (1+\theta t)m(t) \Rightarrow m'(0) = 1$$

$$m''(t) = \theta m(t) + (1+\theta t)m'(t) \Rightarrow m''(0) = \theta + 1$$

$$m'''(t) = 2\theta m'(t) + (1+\theta t)m''(t) \Rightarrow m'''(0) = 3\theta + 1$$

$$m''''(t) = 3\theta m''(t) + (1+\theta t)m'''(t) \Rightarrow m''''(0) = 3\theta^2 + 6\theta +$$

- $$MSE(\hat{\theta}_4) = \frac{1}{n}V(X_1^2) = \frac{1}{n}\left[E(X_1^4) - [E(X_1^2)]^2\right] = \frac{1}{n}\left[3\theta^2 + 6\theta + 1 - (\theta+1)^2\right]$$

$$= \frac{1}{n}\left[2\theta^2 + 4\theta\right] \quad (+3)$$

$\therefore \hat{\theta}_1$ is best. (+1)

(c) (5pts) Which estimators are consistent? Explain why.

All are consistent estimator.

Because $Var(\theta_i) \xrightarrow{n\to\infty} 0$ $(i=1,2,3,4)$ (+5)

< majority of dedection factors >

- only unbiased estimator (−2)
- select only a few out of 4 (−3)

(d) (5pts) Which estimator is MVUE? Justify your answer.

Based on (b) $\hat{\theta}_3$ is MVUE. (+5)

or using sufficient statistic & minimum variance (+5)

(e) (5pts) Find a 95% confidence interval for θ based on $\hat{\theta}_3$.

$$\frac{n\hat{\theta}_3}{\theta} \sim \chi^2(n) \quad (+3)$$

$\Leftrightarrow$ 95% C.I for θ is given by

$$\left(\frac{n\hat{\theta}_3}{\chi^2_{0.975}(n)}, \frac{n\hat{\theta}_3}{\chi^2_{0.025}(n)} \right) \quad (+2)$$

Distribution	PDF or PMF	Mean	Variance	Moment-Generating Function
Continuous distributions				
Uniform	$f(y) = \dfrac{1}{\theta_2 - \theta_1}$	$\dfrac{\theta_1 + \theta_2}{2}$	$\dfrac{(\theta_2 - \theta_1)^2}{12}$	$\dfrac{e^{t\theta_2} - e^{t\theta_1}}{t\,(\theta_2 - \theta_1)}$
Normal	$f(y) = \dfrac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left[-\left(\dfrac{1}{2\sigma^2}\right)(y-\mu)^2\right]$	μ	σ^2	$\exp\left(\mu t + \dfrac{t^2\sigma^2}{2}\right)$
Exponential	$f(y) = \frac{1}{\beta} e^{-y/\beta}$	β	β^2	$\left(1 - \beta t\right)^{-1}$
Gamma	$f(y) = \dfrac{1}{\Gamma(\alpha)\beta^\alpha} y^{\alpha-1} e^{-y/\beta}$	$\alpha\beta$	$\alpha\beta^2$	$\left(1 - \beta t\right)^{-\alpha}$
Chi-square	$f(y) = \dfrac{(y)^{(\nu/2)-1} e^{-y/2}}{2^{\nu/2}\Gamma(\nu/2)}$	ν	2ν	$(1-2t)^{-\nu/2}$
Beta	$f(y) = \left[\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\right] y^{\alpha-1}(1-y)^{\beta-1}$	$\dfrac{\alpha}{\alpha+\beta}$	$\dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	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	$p(y) = p(1-p)^{y-1}$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$	$\dfrac{pe^t}{1-(1-p)e^t}$
Hypergeometric	$p(y) = \dfrac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}}$	$\dfrac{nr}{N}$	$n\left(\dfrac{r}{N}\right)\left(\dfrac{N-r}{N}\right)\left(\dfrac{N-n}{N-1}\right)$	
Poisson	$p(y) = \dfrac{\lambda^y e^{-\lambda}}{y!}$	λ	λ	$\exp[\lambda(e^t - 1)]$
Negative binomial	$p(y) = \binom{y-1}{r-1} p^r (1-p)^{y-r}$	$\dfrac{r}{p}$	$\dfrac{r(1-p)}{p^2}$	$\left[\dfrac{pe^t}{1-(1-p)e^t}\right]^r$
Multinomial	$p(y_1,\ldots,y_k) = \frac{n!}{y_1!y_2!\cdots y_k!} p_1^{y_1}\cdots p_k^{y_k}$			

Standard Normal Probabilities

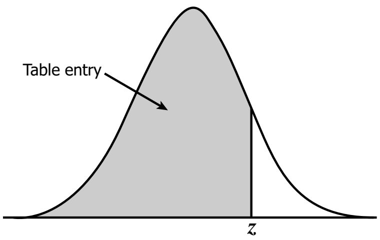

Table entry for z is the area under the standard normal curve to the left of z.

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