Sutomo Salim (CEE / Yr 2)

CV2001 2010/2011 SEM 1

Date No.

1 a) $A = \{1,2,3\}$ $\bar{A} = \{4,5,6\}$

$B = \{1,3,6\}$ $\bar{B} = \{2,4,5\}$

(i) $P(B) = P(1) + P(3) + P(6)$

$= 0.1 + 0.1 + 0.1$

$= 0.3$

(ii) $P(A \cap B) = P(1) + P(3)$

$= 0.1 + 0.1$

$= 0.2$

(iii) $P(\bar{A} | B) = \frac{P(\bar{A} \cap B)}{P(B)}$

$= \frac{P(6)}{0.3}$

$= \frac{1}{3}$

(iv) $P(A | \overline{A \cap B}) = P(A | \bar{A} \cup \bar{B}) = \frac{P(A \cap (\bar{A} \cup \bar{B}))}{P(\overline{A \cap B})}$

$= \frac{P((A \cap \bar{A}) \cup (A \cap \bar{B}))}{1 - P(A \cap B)}$

$= \frac{0 + P(2)}{1 - 0.2}$

$= \frac{1}{4}$

(v) $P(A) = P(1) + P(2) + P(3)$

$= 0.1 + 0.2 + 0.1$

$= 0.4$

$P(A \cap B) = 0.2$

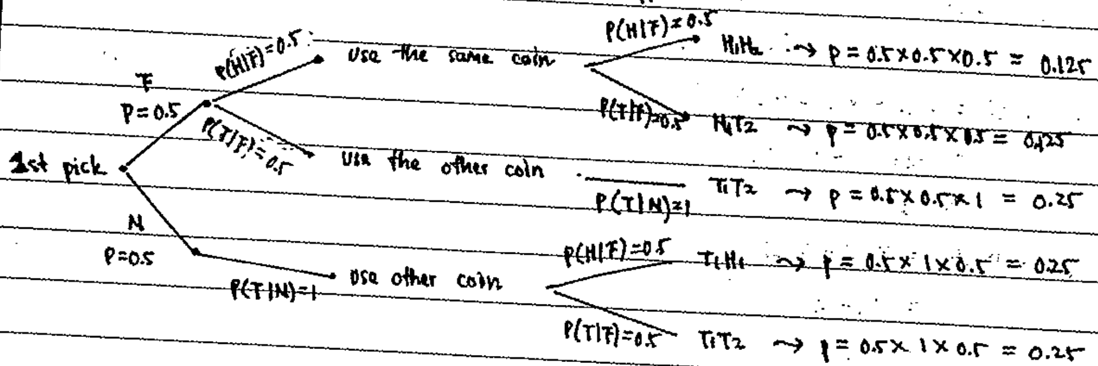
$P(A) \times P(B) = 0.4 \times 0.3 = 0.12$

Since $P(A \cap B) \neq P(A) \times P(B)$; event A and B are not statistically independant

Date No.

b) F = fair coin is picked H = head appears

N = non-fair coin is picked T = tail appears



(i) Prob tail appears on second toss $= P(H_1T_2) + P(T_1T_2) +$

$= 0.125 + (0.25 + 0.25)$

$= \underline{0.625}$ ↳ say $P(A)$

(ii) Prob tail appears on first toss $= P(B) = P(T_1T_2) + P(T_1H_1)$

$= (0.25 + 0.25) + 0.25$

$= 0.75$

this part not necessary

$\therefore P(B \mid A) = \dfrac{P(B \cap A)}{P(A)}$

$= \dfrac{P(T_1T_2)}{0.625}$

$= \dfrac{(0.25 + 0.25)}{0.625}$

$= 0.8$

Date No.

2 a) $f(x) = \frac{dF(x)}{dx} = 2.5 - 0.5x$ when $3 \leq x \leq 5$

$\therefore \quad f(x) = \begin{cases} 2.5 - 0.5x, & 3 \leq x \leq 5 \\ 0, & \text{otherwise} \end{cases}$ $\longrightarrow$ derivative of constants is zero

b) Median requirements $\longrightarrow F(\tilde{x}) = 0.5$

$$\int_3^{\tilde{x}} f(x)\,dx = 0.5$$

$$\left[2.5x - \frac{0.5}{2}x^2\right]_3^{\tilde{x}} = 0.5$$

$$2.5\tilde{x} - 0.25\tilde{x}^2 - 7.5 + 2.25 = 0.5$$

solved $\rightsquigarrow$ $\tilde{x} = 3.5858$ // or $\tilde{x} = 6.4142$ (Not Valid)

Mode $\rightsquigarrow$ find highest value of $f(x)$

Since $f(x)$ equation is linear with negative slope, $f(x)$ max when $x = 3$

$\therefore$ Mode $\doteq x_m = 3$

c) $P(x > x_d) = 0.05$

(i) $P(x < x_d) = 1 - P(x > x_d)$

$= 1 - 0.05$

$= 0.95$

$P(x < x_d) = F(x_d) = 0.95$

$2.5x_d - 0.25x_d^2 - 5.25 = 0.95$

solved $\rightsquigarrow$ $x_d = 4.5528$ // or $x_d = 5.4472$ (Not Valid)

Date No.

(ii) Prob of roof failure occur for the first time in third year ~ Geometric Distribution,

$p = 0.05$

$t = 3$

$$P(t=3) = p\,q^{3-1}$$
$$= (0.05)(0.95)^2$$
$$\approx 0.0451$$

(iii) No penalty within 5 years ~ Binomial Distribution, $p = 0.05$, $n=5$, $x=0$

$$P(x=0) = C^5_0\, p^0 q^{5-0}$$
$$= (0.95)^5$$
$$= 0.7738$$

3 a) $E[z] = \sigma^2$

$$\Rightarrow E\left[a\sum_{k=1}^{n}(X_k - X_{n+k})^2\right]$$
$$= E\left[a\sum_{k=1}^{n}(X_k^2 - 2X_kX_{n+k} + X_{n+k}^2)\right]$$
$$= a\sum_{k=1}^{n}(E[X_k^2] - 2E[X_kX_{n+k}] + E[X_{n+k}^2])$$
$$= a\sum_{k=1}^{n}(\mu^2+\sigma^2 - 2\mu^2 + \mu^2+\sigma^2)$$
$$= a\,n\,2\sigma^2$$
$$\Rightarrow E[z] = \sigma^2 = 2an\sigma^2$$
$$\therefore a = \frac{1}{2n}$$

b) $X_k - X_{n+k} = Y_k$

$$\mathrm{var}[z] = \mathrm{var}\left[\frac{1}{2n}\sum_{k=1}^{n} Y_k\right]$$
$$= \frac{1}{4n^2}\, n\sigma^2$$
$$= \frac{\sigma^2}{4n}$$

$$\lim_{n\to\infty} \frac{\sigma^2}{4n} = 0$$

$\therefore$ consistent

Date No.

3 b) i) unknown μ, unknown σ^2 $\Rightarrow$ use T-distribution to find CI for μ

$$\bar{x} = \frac{492+476+501+467+489+462+509+515+458+475}{10} = 484.4$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{10}(x_i-\bar{x})^2 = 395.16$$

$S = 19.88$

$$-t_{\frac{1-0.95}{2},10-1} \le T \le t_{\frac{1-0.95}{2},10-1}$$

$$-2.262 \le \frac{\bar{x}-\mu}{s/\sqrt{n}} \le 2.262$$

$$470.18 \le \mu \le 498.62$$

ii) $\sigma^2 = \frac{(n-1)s^2}{V}$

$$\chi^2_{1-0.025,9} \le V \le \chi^2_{0.025,9}$$

$$2.7 \le V \le 19.023$$

$$\frac{(n-1)s^2}{19.023} \le \sigma^2 \le \frac{(n-1)s^2}{2.7}$$

$$186.95 \le \sigma^2 \le 1317.2$$

iii) Find n such that $P(|\bar{x}-\mu| \le \varepsilon) = 95\%$

$$\Rightarrow P\left(-\frac{\varepsilon}{s/\sqrt{n}} \le \frac{\bar{x}-\mu}{s/\sqrt{n}} \le \frac{\varepsilon}{s/\sqrt{n}}\right) = 95\%$$

consider 95% CI for T: $P(-T_{0.025,9} \le T \le T_{0.025,9}) = 95\%$

$$\frac{\varepsilon}{s/\sqrt{n}} = T_{0.025,9} = 2.262$$

$$n = \left(\frac{2.262 \times 19.88}{\varepsilon}\right)^2 \rightarrow n \ge 80.89$$

$\varepsilon \le 5$

$\therefore$ Additional 71 specimens should be tested

Date No.

4 a) i) Define r.v X such that the watching rate is $p = P(X=1)$. For the survey results

$$X_i = \begin{cases} 1 & \text{If watch the program} \\ 0 & \text{If not} \end{cases} \quad \text{and} \quad \sum_{i=1}^{100} X_i = 70$$

Since n is big we can use Z-distribution: $\bar{X} = \frac{1}{n}\sum_{i=1}^{100} X_i \sim N(p, \frac{var[X]}{n})$

$$\frac{\bar{X}-p}{\sqrt{var[X]/n}} \sim N(0,1)$$

$\bar{x} = 0.7$

$$var[X] \approx s^2 = \frac{1}{n-1}\left[\sum x_i^2 - n\bar{x}^2\right] = \frac{1}{n-1}\left[\sum x_i - n\bar{x}^2\right]$$

$$= \frac{1}{99}\left[70 - 100(0.7)^2\right]$$

$$= \frac{7}{33}$$

Test $H_0: \mu \geq 0.6$ against $H_1: \mu < 0.6$

$\Rightarrow$ If $Z_0 \leq -Z_\alpha$ we can reject H_0 at significance level $\alpha = 5\%$.

$$Z_0 = \frac{0.7-0.6}{\sqrt{7/33\times100}} \approx 2.171 \quad > \quad -Z_{0.05} = 1.645$$

$\therefore$ Since we cannot reject H_0, we can claim that watching rate of the TV program is over 60% by default

Date No.

ii)

$$X_i = \begin{cases} 1 & \text{if watch the program} \\ 0 & \text{if not} \end{cases} \quad \text{and} \quad \sum_{i=1}^{5} X_i = 4$$

Since n is small, we can only use T-distribution :

$$\bar{x} = \frac{4}{5} = 0.8$$

$$s^2 = \frac{1}{5-1}\left[4-5(0.8)^2\right] = 0.2$$

Test $H_0 : \mu \geq \overset{\mu_0}{\boxed{0.6}}$ against $H_1 : \mu < 0.6$

→ If $T_0 \leq -T_{\alpha,4}$ we can reject H_0 at significance level $\alpha = 5\%$

$$T_0 = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} = \frac{0.8-0.6}{\sqrt{0.2/5}} = 1 \quad > \quad -T_{0.05,4} = -2.132$$

∴ Since we cannot reject H_0, we can claim that the watching rate is over 60%

Date No.

b) i) $Y = \hat{a}_0 + \hat{a}_1 X$

Y	X	X^2	Y^2	XY
2	1.0	1	4	2
3	2.0	4	9	6
3.5	2.5	6.25	12.25	8.75
4.2	3.0	9	17.64	12.6
5	3.5	12.25	25	17.5
5.4	4.0	16	29.16	21.6

$\bar{Y} = 3.85$

$\bar{X} = \frac{8}{3}$

$S_{xx} = \sum X_i^2 - n\bar{x}^2 = 5.833$

$S_{xy} = \sum XY - n\bar{X}\bar{Y} = 6.85$

$\Rightarrow \hat{a}_1 = \frac{S_{xy}}{S_{xx}} = 1.1743$ $\quad \hat{a}_0 = \bar{Y} - \hat{a}_1\bar{X} = 0.7186$

$\therefore Y = 0.7186 + 1.1743X$

ii) Prediction $X_{n+1} = 3.2$

90% CI for Y_{n+1} : $Y_{n+1} = \boxed{\hat{Y}_{n+1}} \pm \boxed{t_{0.05,4}} \sqrt{\frac{SSE}{n-2}} \sqrt{1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{S_{xx}}}$

$\hat{Y}_{n+1} = 0.7186 + 1.1743\ (3.2)$
$= 4.4764$

2.132

$SSE = S_{YY} - \hat{a}_1 S_{xy}$
$= \left(\sum Y^2 - n\bar{Y}^2\right) - 1.1743 \times 6.85$
$= 0.071045$

$\Rightarrow$ solve $\quad Y_{n+1} = 4.4764 \pm 0.3132$

90% CI for Y_{n+1} is $(4.1632, 4.7896)$