PAGE No
DATE / /

P&S2 Assignment

1.

2. Sample mean: $\bar{X} = \frac{\Sigma X_i}{n}$

$E(\Sigma X_i) = \Sigma E(X_i) = \Sigma \mu = n\mu$ (identically distributed)

$V(\Sigma X_i) = \Sigma V(X_i) = n\sigma^2$ (iid)

$E(\bar{X}) = \mu$

$V(\bar{x}) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$

ii Sample variance: $S^2 = \frac{1}{n-1}[\Sigma X_i^2 - n\bar{X}^2]$

$$E(S^2) = \frac{1}{n-1}\left(\Sigma E(X_i^2) - n E(\bar{X}^2)\right)$$

$$= \frac{1}{n-1}\left(\Sigma(\sigma^2 + \mu^2) - n\left(\frac{\sigma^2}{n} + \mu^2\right)\right)$$

$$= \frac{1}{n-1}\left(n(\sigma^2 + \mu^2) - \sigma^2 - n\mu^2\right)$$

$$= \frac{1}{n-1}\left((n-1)\sigma^2\right)$$

$$= \sigma^2$$

iii $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

$V(\chi^2_k) = 2k$

$\therefore V\left(\frac{(n-1)S^2}{\sigma^2}\right) = 2(n-1)$

$$V(S^2) = \frac{\sigma^4}{(n-1)^2} 2(n-1) = \frac{2\sigma^4}{n-1}$$

iv $\sigma^2 = 100 \quad n = 10$

$P(50 < S^2 < 150)$

$Y = \frac{9S^2}{100}$

$Y = \frac{9(50)}{100}$ $\quad (S^2 = 50)$

$= 4.5$

$Y = \frac{9(150)}{100}$ $\quad (S^2 = 150)$

$= 13.5$

$\therefore P(50 < S^2 < 150) = P(4.5 < Y < 13.5)$

$= P(Y < 13.5) - P(Y < 4.5)$

$= 0.8587 - 0.1245$

$= 0.7342$

3.

0	1	2	3	4
86	75	16	2	1

$L = [(1-p)^4]^{86} [4p(1-p)^3]^{75} [6p^2(1-p)^2]^{16} [4p^3(1-p)]^2 [p$ cont

$= [(1-p)^{344}] [p^{75}(1-p)^{225}] [p^{32}(1-p)^{32}] [p^6(1-p)^2] [p]$ cont

$= (1-p)^{603} p^{117}$

$\frac{d \ln L}{dp} = \frac{-603}{(1-p)} + \frac{117}{p} = 0$

$p = \frac{117}{603+117} = \frac{117}{720} = 0.1625$

$\frac{d^2 \ln L}{dp^2} = \frac{-603}{(1-p)^2} - \frac{117}{p^2} < 0$

ii H_0: probabilities conform to Bin(4, p)
H_1: probabilities dont conform to Bin(4, p)
$\tilde{p} = 0.1625$

$P(x=0) = (1-p)^4 = 0.49197$
$P(x=1) = 4p(1-p)^3 = 0.3883$
$P(x=2) = 6p^2(1-p)^2 = 0.11113$
$P(x=3) = 4p^3(1-p) = 0.01437$
$P(x=4) = p^4 = 0.0007$

E_i: 88.55, 68.73, 20, 2.59, 0.13

	O_i	E_i	$(O_i - E_i)/E_i$
0	86	88.55	0.074
1	75	68.73	0.572
2+	19	22.72	0.608
			1.254 ← χ^2_{cal}

$DOF = 3-1-1 = 1$

$\chi^2_{cal} < \chi^2_{tab}$ for 5%.
∴ Model is a good fit

4. $f(x) = \dfrac{c\beta^3}{(x+\beta)^4}$ $x>0, \beta>0$

$f(x) = \dfrac{\alpha\lambda^\alpha}{(\lambda+x)^{\alpha+1}}$ $x>0$ $\lambda>0$ $\alpha>0$ (table book)

$E(x) = \dfrac{\lambda}{(\alpha-1)}$ $V(x) = \dfrac{\alpha\lambda^2}{(\alpha-1)^2(\alpha-2)}$ $\alpha>2$

∴ above pdf can be identified with $c=3$, $\lambda=\beta$
∴ $E(x) = \dfrac{\beta}{2}$ $V(x) = \dfrac{3}{4}\beta^2$

$$\int_0^\infty \frac{c\beta^3}{(x+\beta)^4}\,dx = 1$$

$$\left[\frac{-c}{3}\frac{\beta^3}{(x+\beta)^3}\right]_0^\infty = 1$$

$$c = 3$$

ii Using MOM:

$$\bar{X}_n = \frac{\beta}{2}$$

$$\tilde{\beta} = 2\bar{X}_n$$

$$MSE[\tilde{\beta}] = V[\tilde{\beta}] + (Bias[\tilde{\beta}])^2$$
$$E(\tilde{\beta}) = E(2\bar{X}_n) = 2E(\bar{X}_n) = 2E(x) = 2(\beta/2) = \beta$$
$$Bias = E(\tilde{\beta}) - \beta = 0$$
$\therefore$ Unbiased

$$V(\tilde{\beta}) = V(2\bar{X}_n) = 4V(\bar{X}_n) = \frac{4V(x)}{n} = \frac{4\left(\frac{3}{4}\beta^2\right)}{n} = \frac{3\beta^2}{n}$$

$$\therefore MSE(\tilde{\beta}) = \frac{3\beta^2}{n} + 0 = \frac{3\beta^2}{n}$$

$\therefore$ Consistent since MSE tends to 0 as $n \to \infty$

iii $$MSE(b\bar{X}_n) = V(b\bar{X}_n) + (Bias(b\bar{X}_n))^2$$
$$= b^2\frac{3}{4}\frac{\beta^2}{n} + (E(b\bar{X}_n) - \beta)^2$$
$$= b^2\frac{3}{4}\frac{\beta^2}{n} + \left(\beta\left(\frac{b}{2} - 1\right)\right)^2$$
$$= \frac{\beta^2}{n}\left(\frac{3}{4}b^2 + n\left(\frac{b}{2} - 1\right)^2\right)$$

$$\frac{d\,MSE}{db} = \frac{\beta^2}{n}\left(\frac{3}{2}b + n\left(\frac{b}{2} - 1\right)\right) = 0$$

PAGE No
DATE / /

5.i $E(x) = \frac{a+b}{2}$

$b = 2E(x) - a$
$= 2\bar{x} - a$

$V(x) = \frac{(b-a)^2}{12}$

$s^2 = \frac{(2\bar{x} - a - a)^2}{12} = \frac{(\bar{x} - a)^2}{3}$ $\qquad (\bar{x} - a)^2 = 3s^2$

$\tilde{a} = \bar{x} - \sqrt{3}s$

ii $\tilde{b} = 2\bar{x} - (\bar{x} - \sqrt{3}s)$
$= \bar{x} + \sqrt{3}\, s$

iii Sample : 1, 2, 3, 4, 50
$\bar{x} = 12, \quad s = 21.27$
MOM estimates:
$\tilde{a} = 12 - \sqrt{3}(21.27)$ $\qquad \tilde{b} = 12 + \sqrt{3}(21.27)$
$= -24.84$ $\qquad = 48.84$

For $U(a, b)$, probability of sample point being $< a$ or $> b$
& we have sample value 50 which is $> b$.

iv $L(b) = \frac{1}{b^n}$ if $b \geq \max(x_i)$

otherwise $L = 0$
Diff w.r.t b doesn't work because in range of x depends on b
We must find b that maximizes $L(b)$ for $\max(x_i)$
We want b to be as small as possible subject to constraint
that $b \geq \max(x_i)$
Clearly max. attained at $b = \max(x_i)$
$\therefore \hat{b} = \max(x_i)$

6. $P(\text{Type I error}) = P(X \geq 3 \mid p=0.1) + P(X=0 \mid p=0.1)\,P(Y \geq 5 \mid p=0.1)$
$+ P(X=1 \mid p=0.1)\,(P(Y \geq 4 \mid p=0.1) + P(X=2 \mid p=0.1)$
$P(Y \geq 3 \mid p=0.1)$
$= (1-0.8891) + (0.2824)(1-0.9957) +$
$(0.659-0.2824)(1-0.9744) + (0.8891-0.659)$
$(1-0.8891)$
$= 0.14727 \approx 15\%$

ii Prob. of rejecting H_0 when $p=0.3$
$= P(X \geq 3 \mid p=0.3) + P(X=0 \mid p=0.3)\,P(Y \geq 5 \mid p=0.3) + P(X=1 \mid p=0.3)$
$P(Y \geq 4 \mid p=0.3) + P(X=2 \mid p=0.3)\,P(Y \geq 3 \mid p=0.3)$
$= (1-0.2528) + (0.0138)(1-0.7237) + (0.085-0.0138)$
$(1-0.6925) + (0.2528-0.085)(1-0.2528)$
$= 0.91253 \approx 91\%$

iii $P(\text{type II error}) = 1-0.91253 = 0.08747 \approx 9\%$

iv $n=120,\ X=40$ $\qquad \hat{p} = \frac{40}{120} = \frac{1}{3} = 0.333$

$$\frac{\frac{X}{n} - p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0,1)$$

$Z_{10\%} = 1.2816$

$$\hat{p} - 1.2816\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.333 - 1.2816\sqrt{\frac{0.333(1-0.333)}{120}}$$

$= 0.2782$

v $H_0: p = 0.278 \qquad H_1: p > 0.278$

$$\frac{X - np_0}{\sqrt{np_0q_0}} \sim N(0,1)$$

$$\frac{39.5 - 120(0.278)}{\sqrt{120(0.278)(0.722)}} = 1.2511 \cdot Z_{cal}$$

PAGE No
DATE / /

$Z_{cal} < Z_{tab}$

$\therefore$ Do not reject H_0

vi Lower bound of CI implies there's only 10% chance of $p \leq 0.278$
Hypothesis test says p could be ≤ 0.278 with prob. > 10%

vii $H_0 : \mu_1 = \mu_2$ $\quad H_1 : \mu_1 \neq \mu_2$

$$\frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1 + n_2 - 2}$$

$\bar{X}_1 = 65$ $\quad \bar{X}_2 = 70$ $\quad S_1 = 54$ $\quad S_2 = 70$ $\quad n_1 = 12$ $\quad n_2 = 15$

$$S_p^2 = \frac{1}{25}\left(11(2916) + 14(4900)\right)$$
$$= 4027.04$$

$$\frac{65 - 70 - 0}{63.459\sqrt{\frac{1}{12} + \frac{1}{15}}} = -0.2034$$

Within $\pm 2.06 \, (t_{25, 2.5\%})$

$\therefore$ Do not reject H_0

PAGE No
DATE / /

2.	City 1	City 2	City 3	City 4	City 5	City 6
Pub	24	45	29	33	20	[illegible]
Pvt.	30	54.5	30	40	28.5	3[illegible]

	City 7	City 8	City 9
Pub	26.5	25	27.5
Pvt	30.5	30.5	35.5

i $\Sigma X_{Pub} = 270$

$(\bar{x}_{Pub}) = 30$

$\Sigma X_{Pvt} = 315.5$

$(\bar{x}_{Pvt}) = 35.06$

$\Sigma X^2_{Pub} = 8614.5$

$$S^2_{Pub} = \frac{(8614.5 - 9 \times 30^2)}{8} = 64.31$$

$\Sigma X^2_{Pvt} = 11599.25$

$$S_{Pvt}^2 = \frac{(11599.25 - 9 \times 35.06^2)}{8} = 67.42$$

ii 2 sided t-test in case samples are from populations of equal variances

$H_0 : \sigma_1^2 = \sigma_2^2$

$H_1 : \sigma_1^2 \neq \sigma_2^2$

$$T.S = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \sim F_{n_1-1, n_2-1}$$

$$= \frac{64.31}{67.42} = 0.9542$$

$F_{8,8}$ at 5% LOS = 0.2256, 4.432

$\therefore$ Do not Reject H_0

iii $H_0: \mu_1 = \mu_2$

$H_1: \mu_1 < \mu_2$

$$T.S = \frac{(\bar{X}_2 - \bar{X}_1) - (\mu_2 - \mu_1)}{\sqrt{S_p^2 (1/n_1 + 1/n_2)}} \sim t_{n_1+n_2-2}$$

$$S_p^2 = \frac{S_1^2(n_1-1) + S_2^2(n_2-1)}{n_1+n_2-2}$$

$$= \frac{64.31 \times 8 + 67.42 \times 8}{16} = 65.86$$

$$T.S = \frac{(35.036 - 30) - 0}{\sqrt{65.86(1/9 + 1/9)}} = 1.32$$

$$P(t_{16} > 1.32) = 20.5\%$$

8

i When $E(\hat{\theta}) = \theta$

ii $E(\hat{\theta}) - \theta$

iii $(E(\hat{\theta}) - \theta)^2$

iv The one with lower MSE

v If MSE converges to 0 as $n \to \infty$

vi

a MOM: population moments equated to sample moments to estimate parameters

b MLE: Max. likelihood functions generated. MLE of parameter given by soln. to $\frac{dL}{d\theta} = 0$.

PAGE No
DATE / /

9.

i $t_k = \frac{N(0,1)}{\sqrt{\frac{\chi^2_k}{K}}}$ K = DOF

ii Mean = 0
Variance = $\frac{K}{K-2}$

iii a $\frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$
$t_9 = 3.25$

$CI = 50 - 3.25 \times \sqrt{\frac{48.667}{10}}, \; 50 + 3.25 \times \sqrt{\frac{48.667}{10}}$
$= (42.83, 57.17)$

b $t_{n-1} \sim N\left(0, \sqrt{\frac{n-1}{n-3}}\right) \sim N\left(0, \sqrt{\frac{9}{7}}\right)$
$\frac{\bar{X} - \mu}{S/\sqrt{n}} \sim N\left(0, \sqrt{\frac{9}{7}}\right)$
$\frac{X - \mu}{S/\sqrt{7n/9}} \sim N(0,1)$
Critical val = 2.58
$CI = \left(50 \pm 2.58 \times \sqrt{\frac{48.667}{10} \times \frac{9}{7}}\right)$
$= (43.54, 56.45)$

10.

i Event of rejecting hypotheses when it's true.

$X \sim Poi(n\mu) \sim Poi(3000)$

$H_0 : X \sim Poi(3000)$

$P(\text{reject } H_0 \text{ when } H_0 \text{ true}) = P(x > 3100 \text{ when } x \sim Poi(3000))$

Using normal approximation

$X \sim N(3000, 3000)$

$P(x > 3100) = P\left(z > \frac{3100 - 3000}{\sqrt{3000}}\right) = 1 - P(z < 1.826) = 0.0339$

ii Event of accepting hypotheses when it's false

iii Prob. of rejecting hypotheses when false

$P(\text{reject } H_0 \text{ when false}) = P(x > 3100 \text{ when } x \sim Poi(n\mu) \sim N(n\mu, n\mu))$

$P(x > 3100) = 1 - \left(z < \frac{3100 - n\mu}{\sqrt{n\mu}}\right)$

iv $\hat{\mu} \sim N\left(\mu, \frac{\hat{\mu}}{n}\right)$

$\frac{\hat{\mu} - \mu}{\sqrt{\hat{\mu}/n}} \sim N(0,1)$

$P\left(-2.5758 < \frac{2.9 - \mu}{\sqrt{2.9/1000}} < 2.5758\right) = 0.99$

$\therefore C.I = (2.7613, 3.0387)$

11. Sample mean (Revised) $= \frac{213200 - 11000 - 48000 + 27500 + 31500}{12}$

$= 17767$

Σx^2 (Revised) $= 4919860000 - 121000000 - 2304000000$
$+ 756250000 + 992250000$
$= 4243360000$

$\therefore$ Sample sd (Revised) $= \left(\frac{1}{11}\left(4243360000 - \frac{213200^2}{12}\right)\right)^{1/2}$

$= \left(\frac{1}{11}(455506166.7)\right)^{1/2}$

$= 41409696.97^{1/2}$

$= 6435$

No change in sample mean but a reduction in sample sd.

Mean unaltered as total salary of temporary employees being replaced = total salary of permanent employees who replaced them. Reduction in s.d due to salaries of permanent employees closer to sample mean.

12.

i CRLB doesn't hold true where range of distribution involves parameter. This is due to discontinuity, so derivative in formula doesn't make sense.

ii ~~$Y = \max_i X_i$~~
~~For $0 < y < \theta$~~
~~$P(Y < y) = P[\max_i X_i < y] = P[\bigcap_{i=1}^{n}$~~

iii $Bias[\hat{\theta}(c)] = E[\hat{\theta}(c) - \theta]$

$= E[cY - \theta] = c\,E[Y] - \theta$

$= \frac{cn\theta}{n+1} - \theta$

$= \left(\frac{cn}{n+1} - 1\right)\theta$

$MSE(\hat{\theta}(c)) = E[(\hat{\theta}(c) - \theta)^2]$

$= E[(cY - \theta)^2]$

$= E[c^2Y^2 - 2c\theta Y + \theta^2]$

$= c^2 E(Y^2) - 2c\theta E(Y) + \theta^2$

$= c^2 \frac{n\theta^2}{n+2} - c\frac{2n\theta^2}{n+1} + \theta^2$

iv $Bias[\hat{\theta}(c_u)] = 0$

$\left(\frac{c_u n}{n+1} - 1\right)\theta = 0$

$c_u = \frac{n+1}{n}$

v $\frac{d}{dc} MSE[\hat{\theta}(c)] = 0$

$\frac{d}{dc}\left[c^2 \frac{n\theta^2}{n+2} - c\frac{2n\theta^2}{n+1} + \theta^2\right] = 0$

$2c\frac{n\theta^2}{n+2} - \frac{2n\theta^2}{n+1} = 0$

$c_m = \frac{2n\theta^2}{n+1} \cdot \frac{n+2}{2n\theta^2} = \frac{n+2}{n+1}$

vi $\hat{\theta}(c_m)$ due to lower MSE

As n becomes large, $\hat{\theta}(c_u) = \frac{n+1}{n} Y \to Y$,

$\hat{\theta}(c_m) = \frac{n+2}{n+1} Y \to Y$