1. (a)

| | $\theta_1$ | $\theta_2$ | $\theta_3$ |
|---|---|---|---|
| $a_1$ | 2 | 1 | 5 |
| $a_2$ | 0 | 4 | 6 |
| $a_3$ | 4 | 0 | 2 |

| | $\theta_1$ | $\theta_2$ | $\theta_3$ |
|---|---|---|---|
| $x_1$ | 0.4 | 0.5 | 0.3 |
| $x_2$ | 0.6 | 0.5 | 0.7 |

Let

| | $x_1$ | $x_2$ |
|---|---|---|
| $d_1$ | $a_1$ | $a_1$ |
| $d_2$ | $a_1$ | $a_2$ |
| $d_3$ | $a_2$ | $a_3$ |

then since $R(\theta, d) = E(L(\theta, d)) = \sum L(\theta, d) P(x|\theta)$

$R(\theta_1, d_1) = 2 \cdot 0.4 + 2 \cdot 0.6 = 2$

so

$R(\theta_2, d_1) = 1 \cdot 0.5 + 1 \cdot 0.5 = 1$

$R(\theta_3, d_1) = 5 \cdot 0.3 + 5 \cdot 0.7 = 5$

$R(\theta_1, d_2) = 2 \cdot 0.4 + 0 \cdot 0.6 = 0.8$

$R(\theta_2, d_2) = 1 \cdot 0.5 + 4 \cdot 0.5 = 2.5$

$R(\theta_3, d_2) = 5 \cdot 0.3 + 6 \cdot 0.7 = 5.7$

$R(\theta_1, d_3) = 0 \cdot 0.4 + 4 \cdot 0.6 = 2.4$

$R(\theta_2, d_3) = 4 \cdot 0.5 + 0 \cdot 0.5 = 2$

$R(\theta_3, d_3) = 6 \cdot 0.3 + 2 \cdot 0.7 = 3.2$

(b) $\max_\theta R(\theta, d_1) = 5$

$\max_\theta R(\theta, d_2) = 5.7$

$\max_\theta R(\theta, d_3) = 3.2$

$\Rightarrow \min(5, 5.7, 3.2) = 3.2 \Rightarrow$ minimax : $d_3$

(c) $\pi(\theta_1) = \pi(\theta_2) = \pi(\theta_3) = \frac{1}{3}$

$B(d) = E(R(\theta, d)) = \sum R(\theta_i, d) \pi(\theta_i)$

$B(d_1) = (2+1+5) \cdot \frac{1}{3} = \frac{8}{3}$

$B(d_2) = (0.8+2.5+5.7) \cdot \frac{1}{3} = 3$

$B(d_3) = (2.4+2+3.2) \cdot \frac{1}{3} = \frac{7.6}{3}$

$\Rightarrow B(d_3)$ attains minimum, Bayes : $d_3$

(d) $P(x_2) = 0.6 \cdot \frac{1}{3} + 0.5 \cdot \frac{1}{3} + 0.7 \cdot \frac{1}{3} = 0.6$

$P(\theta_1 | x_2) = \frac{0.6 \cdot \frac{1}{3}}{0.6} = \frac{1}{3}$, $P(\theta_2 | x_2) = \frac{0.5 \cdot \frac{1}{3}}{0.6} = \frac{5}{18}$, $P(\theta_3 | x_2) = \frac{0.7 \cdot \frac{1}{3}}{0.6} = \frac{7}{18}$

$PR(a_1) = 2 \cdot \frac{1}{3} \cdot 1 \cdot \frac{5}{18} + 5 \cdot \frac{7}{18} = \frac{26}{9}$

$PR(a_2) = 0\cdot\frac{1}{3} + 4\cdot\frac{5}{18} + 6\cdot\frac{7}{18} = \frac{31}{9}$

$PR(a_3) = 4\cdot\frac{1}{3} + 0\cdot\frac{5}{18} + 2\cdot\frac{7}{18} = \frac{19}{9}$

$\Rightarrow$ $PR(a_3)$ attains minimum, Bayes : $a_3$

8. (a) $P(A)+P(B)=1$ and $P(B)=2P(A) \Rightarrow P(A)=\frac{1}{3},\ P(B)=\frac{2}{3}$

$f(x|A) = \frac{1}{\sqrt{2\pi}\sigma}\exp\{-\frac{x^2}{2\sigma^2}\}$

$f(x|B) = \frac{1}{\sqrt{2\pi}\sigma}\exp\{-\frac{(x-1)^2}{2\sigma^2}\}$

under 0-1 loss, Bayes rule is A if

$h(A|x) > h(B|x) \Leftrightarrow f(x|A)P(A) > f(x|B)P(B)$, and B, otherwise

$\Rightarrow \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{x^2}{2\sigma^2})\cdot\frac{1}{3} > \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{(x-1)^2}{2\sigma^2})\cdot\frac{2}{3}$

$\Rightarrow \exp(-\frac{x^2}{2\sigma^2}) > 2\exp(-\frac{(x-1)^2}{2\sigma^2})$

$\Rightarrow -x^2 > 2\sigma^2\ln(2) - (x-1)^2$

$\Rightarrow x < (1-2\sigma^2\ln(2))/2 := c$

so $\begin{cases} \text{classify to A if } x < (1-2\sigma^2\ln(2))/2 \\ \text{classify to B if o/w} \end{cases}$

(b) The proportion of observations that will be classified as A is

$P(c > X) = P(c > X|\theta_1)\pi(\theta_1) + P(c > X|\theta_2)\pi(\theta_2)$

$= P(c > X|\theta_1)\cdot\frac{1}{3} + P(c > X|\theta_2)\cdot\frac{2}{3}$

| $\sigma^2$ | $P(c>X\|\theta_1)$ | $P(c>X\|\theta_2)$ | $P(c>X)$ |
|---|---|---|---|
| 0.1 | 0.91 | 0.03 | 0.32 |
| 1 | 0.42 | 0.11 | 0.22 |
| 25 | 0.00038 | 0.00018 | 0.00025 |
| 100 | $\approx 0$ | $\approx 0$ | $\approx 0$ |

9

| | $H: p \le \frac{1}{2}$ | $K: p > \frac{1}{2}$ |
|---|---|---|
| $a_1$: choose H | 0 | 1 |
| $a_2$: choose K | 2 | 0 |

| X | P(X\|p) |
|---|---|
| 0 | $(1-p)^3$ |
| 1 | $3p(1-p)^2$ |
| 2 | $3p^2(1-p)$ |
| 3 | $p^3$ |

| | X=0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $d_1$ | $a_1$ | $a_2$ | $a_2$ | $a_2$ |
| $d_2$ | $a_1$ | $a_1$ | $a_2$ | $a_2$ |
| $d_3$ | $a_1$ | $a_1$ | $a_1$ | $a_2$ |

$$R(p,d_1) = \begin{cases} 2(1-(1-p)^3), & p \le \frac{1}{2} \\ 1-p^3, & p > \frac{1}{2} \end{cases}$$

$$R(p,d_2) = \begin{cases} 2(3p^2(1-p)+p^3), & p \le \frac{1}{2} \\ (1-p)^3+3p(1-p)^2, & p > \frac{1}{2} \end{cases}$$

$$R(p,d_3) = \begin{cases} 2p^3, & p \le \frac{1}{2} \\ 1-p^3, & p > \frac{1}{2} \end{cases}$$

$$B(d_1) = \int_0^{0.5} 2(1-(1-p^3))\,dp + \int_{0.5}^1 (1-p)^3\,dp = \frac{35}{64}$$

$$B(d_2) = \int_0^{0.5} 2(3p^2(1-p)+p^3)\,dp + \int_{0.5}^1 (1-p)^3+3p(1-p)^2\,dp = \frac{9}{32}$$

$$B(d_3) = \int_0^{0.5} 2p^3\,dp + \int_{0.5}^1 1-p^3\,dp = \frac{19}{64}$$

$\Rightarrow B(d_2)$ attains minimum, Bayes : $d_2$

11. $d_1 : \hat{p} = \frac{X}{2}$, $d_2 : \tilde{p} = \frac{X+1}{3}$, $d_3 : \bar{p} = \frac{X+1}{4}$

$R(p, d) = MSE(d) = Var(d) + (Bias(d))^2$

$Bias(\frac{X}{2}) = E(\frac{X}{2}) - p = 0$

$Var(\frac{X}{2}) = \frac{2p(1-p)}{4} = \frac{p(1-p)}{2}$

$R(\hat{p}, p) = \frac{p-p^2}{2}$

$Bias(\frac{X+1}{3}) = \frac{2p+1}{3} - p = \frac{1-p}{3}$

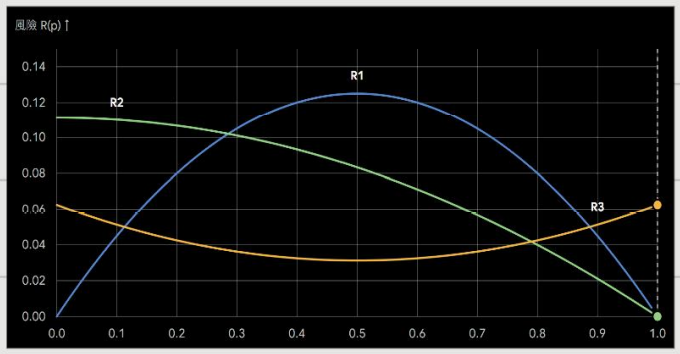
$Var(\frac{X+1}{3}) = \frac{2p(1-p)}{9}$

$R(\tilde{p}, p) = \frac{2p(1-p)}{9} + \frac{(1-p)^2}{9} = \frac{1-p^2}{9}$

$Bias(\frac{X+1}{4}) = \frac{1-2p}{4}$

$Var(\frac{X+1}{4}) = \frac{2p(1-p)}{16}$

$R(\bar{p}, p) = \frac{p(1-p)}{8} + (\frac{1-2p}{4})^2 = \frac{2p^2+1-2p}{16}$




Since each $\hat{p}, \tilde{p}, \bar{p}$ perform better than the other in different interval of $p \in [0,1]$

$\Rightarrow$ No estimate dominated others.

13. $K \sim B(n, p)$, then $\pi(k) = \binom{n}{k} p^k (1-p)^{n-k}$

Let $X = \begin{cases} 1, & \text{抽到瑕疵} \\ 0, & \text{抽到正常} \end{cases}$

$$P(X=1 \mid K=k) = \frac{k}{n}$$

$$P(X=0 \mid K=k) = \frac{n-k}{n}$$

$$P(X=1) = \sum_{k=1}^{n} \frac{k}{n} \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=1}^{n} p\left[\binom{n-1}{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)}\right]$$

$$= p \underbrace{\sum_{y=0}^{n-1} \binom{n-1}{y} p^y (1-p)^{(n-1)-y}}_{\text{pmf of } B(n-1,p)} = p \Rightarrow \text{故 } P(X=0) = 1-P(X=1) = 1-p$$

(a) For $k = 1, 2, \cdots, n$,

$$P(k \mid X=1) = \frac{k}{n} \binom{n}{k} p^k (1-p)^{n-k} / p$$

$$= \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k}$$

(b) For $k = 0, 1, \cdots, n-1$,

$$P(k \mid X=0) = \frac{n-k}{n} \binom{n}{k} p^k (1-p)^{n-k} / (1-p)$$

$$= \binom{n-1}{k} p^k (1-p)^{n-k-1}$$

# 15.

(a)

| | $\theta_1$ | $\theta_2$ | $\theta_3$ |
|---|---|---|---|
| $x_1$ | .1 | .2 | .4 |
| $x_2$ | .1 | .2 | .2 |
| $x_3$ | .2 | .2 | .2 |
| $x_4$ | .6 | .4 | .2 |

Suppose that $x_2$ is observed, the posterior distribution is

$$p(\theta_i \mid x_2) = \frac{p(x_2 \mid \theta_i)\, p(\theta_i)}{\sum_{j=1}^{3} p(x_2 \mid \theta_j)\, p(\theta_j)}$$

$\theta_1 = 1$ : $p(x_2 \mid \theta_1)\, p(\theta_1) = 0.1 \times 0.5 = 0.05$

$\theta_2 = 10$ : $p(x_2 \mid \theta_2)\, p(\theta_2) = 0.2 \times 0.25 = 0.05$ $\Rightarrow$ $p(x_2) = 0.05 + 0.05 + 0.05 = 0.15$

$\theta_3 = 20$ : $p(x_2 \mid \theta_3)\, p(\theta_3) = 0.2 \times 0.25 = 0.05$

$$p(\theta_1 \mid x_2) = \frac{0.05}{0.15} = \frac{1}{3}$$

$$p(\theta_2 \mid x_2) = \frac{0.05}{0.15} = \frac{1}{3}$$

$$p(\theta_3 \mid x_2) = \frac{0.05}{0.15} = \frac{1}{3}$$

$\Rightarrow$ $p(\theta_1) = \frac{1}{3}$, $p(\theta_2) = \frac{1}{3}$, $p(\theta_3) = \frac{1}{3}$

(b) Under squared loss, the Bayes estimate is the mean of posterior distribution of $\Theta$

$$\hat{\theta}_{Bayes} = E(\Theta \mid x_2) = \sum_{i=1}^{3} \theta_i\, p(\theta_i \mid x_2) = 1 \times \frac{1}{3} + 10 \times \frac{1}{3} + 20 \times \frac{1}{3} = \frac{31}{3}$$

(c) For the loss function $\ell(\theta, \hat{\theta}) = |\theta - \hat{\theta}|$, the Bayes estimate is the median of posterior distribution of $\Theta$

$p(\Theta \le 1 \mid x_2) = \frac{1}{3} \approx 0.33$ and $p(\Theta \le 10 \mid x_2) = \frac{1}{3} + \frac{1}{3} \approx 0.67$ $\Rightarrow$ Median $= 10$ $\Rightarrow$ $\hat{\theta}_{Bayes} = 10$.

(d) 以 (a) 小題的後驗分佈作為新的先驗分佈：$p_{new}(\theta_i) = \frac{1}{3}$

$\theta_1$ : $p(x_1 \mid \theta_1)\, p_{new}(\theta_1) = 0.1 \times \frac{1}{3} = \frac{1}{30}$

$\theta_2$ : $p(x_1 \mid \theta_2)\, p_{new}(\theta_2) = 0.2 \times \frac{1}{3} = \frac{2}{30}$ $\Rightarrow$ $p(x_1) = \frac{1}{30} + \frac{2}{30} + \frac{4}{30} = \frac{7}{30}$

$\theta_3$ : $p(x_1 \mid \theta_3)\, p_{new}(\theta_3) = 0.4 \times \frac{1}{3} = \frac{4}{30}$

$$p(\theta_1 \mid x_1) = \frac{1/30}{7/30} = \frac{1}{7}$$

$$p(\theta_2 \mid x_1) = \frac{2/30}{7/30} = \frac{2}{7}$$

$$p(\theta_3 \mid x_1) = \frac{4/30}{7/30} = \frac{4}{7}$$

$\Rightarrow$ $p(\theta_1) = \frac{1}{7}$, $p(\theta_2) = \frac{2}{7}$, $p(\theta_3) = \frac{4}{7}$

#23.

(a) $X_1, \dots, X_9 \mid \mu \overset{iid}{\sim} N(\mu, 9)$, $\mu \sim N(0,1)$

$\pi(\mu \mid \underline{x}) \propto \prod_{i=1}^{9} f(x_i \mid \mu) \cdot \pi(\mu)$

$\propto \exp\left(-\frac{1}{2} \cdot (\cdots)\mu^2 + \bar{x} \cdot \mu\right)$

$\propto \exp\left(-\frac{1}{2} \cdot (\cdots)\left(\mu - \frac{\bar{x}}{2}\right)^2\right) \Rightarrow \mu \mid \underline{x} \sim N\left(\frac{\bar{x}}{2}, \frac{1}{2}\right)$

Mean $= E(\mu \mid \underline{x}) = \frac{\bar{x}}{2}$ is the Bayesian estimator, precision $= \frac{1}{\mathrm{var}(\mu \mid \underline{x})} = 2$.

(b) $\int_{-\infty}^{\infty} E\left[\left(\frac{\bar{x}}{2} - \mu\right)^2 \mid \mu\right] \pi(\mu)\, d\mu$

$= \int_{-\infty}^{\infty} \left[\mathrm{Var}\left(\frac{\bar{x}}{2} - \mu \mid \mu\right) + \left[E\left(\frac{\bar{x}}{2} - \mu \mid \mu\right)\right]^2\right] \pi(\mu)\, d\mu$

$= \frac{1}{4} + \frac{1}{4}\left(\mathrm{var}(\mu) + [E(\mu)]^2\right) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

(c) Using the posterior distribution $\mu \mid \underline{x} \sim N\left(\frac{\bar{x}}{2}, \frac{1}{2}\right)$, a 95% credibility interval of $\mu$ is $\left[\frac{\bar{x}}{2} - z_{0.025}\sqrt{\frac{1}{2}},\ \frac{\bar{x}}{2} + z_{0.025}\sqrt{\frac{1}{2}}\right]$ and $z_{0.025} = 1.96$.

#24

$\mu \sim N(0,1)$, $X \mid \mu \sim N(\mu, 0.25)$

$\pi(\mu \mid x) \propto f(x \mid \mu)\pi(\mu)$

$\propto \exp\left(-\frac{1}{2}\mu^2 - \frac{1}{2 \times 0.25}(x-\mu)^2\right) = \exp\left(-\frac{1}{2}\left(\mu^2 + \frac{(x-\mu)^2}{0.25}\right)\right)$

$= \exp\left(-\frac{1}{2}\left(\mu^2 + \frac{1}{0.25}x^2 - \frac{2}{0.25}x\mu + \frac{1}{0.25}\mu^2\right)\right) = \exp\left(-\frac{1}{2}\left(\frac{5}{4}\mu^2 - \frac{2}{0.25}x\mu + \frac{1}{0.25}x^2\right)\right)$

$\propto \exp\left(-\frac{1}{2}\frac{1}{\frac{4}{5}}\left(\mu - \frac{4}{5}x\right)^2\right)$

$\Rightarrow \mu \mid x \sim N\left(\frac{4}{5}x, \frac{4}{5}\right)$

For squared loss, the Bayes estimator of $\mu$ is $E(\mu \mid x) = \frac{4}{5}x$

MLE: $L(\mu; x) = \frac{1}{\sqrt{2\pi}\,0.5}\exp\left(-\frac{1}{2 \times 0.25}(x-\mu)^2\right)$

$\ell(\mu) = \log\left(\frac{1}{\sqrt{2\pi}\,0.5}\right) - \frac{1}{2 \times 0.25}(x-\mu)^2$

$\frac{\partial \ell}{\partial \mu} \overset{\text{set}}{=} 0 \quad \hat{\mu}_{MLE} = x$

Risk of MLE: $R(\mu, \hat{\mu}_{MLE}) = Var(X|\mu) = \frac{1}{4}$

Risk of Bayes estimate: $R(\mu, \hat{\mu}_{Bayes}) = Var(\frac{4}{5}X|\mu) + Bias(\hat{\mu}_{Bayes})^2$

$Var(\frac{4}{5}X|\mu) = \frac{16}{25} Var(X|\mu) = \frac{4}{25}$

$Bias = E[\frac{4}{5}X|\mu) - \mu] = \frac{4}{5}\mu - \mu = \frac{1}{5}\mu \Rightarrow Bias^2 = \frac{1}{25}\mu^2$

$\Rightarrow R(\mu, \hat{\mu}_{Bayes}) = \frac{4+\mu^2}{25}$

$\frac{4+\mu^2}{25} < \frac{1}{4} \Rightarrow 16 + 4\mu^2 < 25 \Rightarrow |\mu| < \frac{3}{2}$

Ans: 當 $|\mu| < \frac{3}{2}$ 時，貝氏估計量的風險比 MLE 小，當 $|\mu| > \frac{3}{2}$ 則反之

#28.

(a) $X_1, \dots, X_n | \theta \overset{iid}{\sim} Poi(\theta)$ , $\theta \sim P(\alpha, \lambda)$

$\pi(\theta|\underline{x}) \propto \prod_{i=1}^n f(x_i|\theta)\pi(\theta) \propto e^{-\theta n} e^{-\lambda\theta} \theta^{(\alpha + \sum_{i=1}^n x_i - 1)} = e^{-\theta(n+\lambda)} \theta^{(\alpha + n\bar{x} - 1)}$

$\theta|\underline{x} \sim P(\alpha + n\bar{x}, n+\lambda)$ , the posterior distribution is a Gamma.

(b) $\hat{\theta}_{Bayes} = E(\theta|\underline{x}) = \frac{\alpha + n\bar{x}}{n+\lambda} = \frac{n}{n+\lambda}\bar{x} + \frac{\lambda}{n+\lambda} \cdot \frac{\alpha}{\lambda}$ ,

$\frac{\alpha}{\lambda}$ is mean of $P(\alpha, \lambda)$ and $(\frac{n}{n+\lambda}, \frac{\lambda}{n+\lambda})$ are weighes because $\frac{n}{n+\lambda} + \frac{\lambda}{n+\lambda} = 1$

(c) $R(\theta, \frac{\alpha + n\bar{x}}{n+\lambda}) = E[\theta - \frac{\alpha + n\bar{x}}{n+\lambda}]^2 = Var(\theta - \frac{\alpha + n\bar{x}}{n+\lambda}) + [E(\frac{\alpha + n\bar{x}}{n+\lambda} - \theta)]^2$

$= (\frac{n}{n+\lambda})^2 Var(\bar{x}|\theta) + [E(\frac{\alpha + n\bar{x}}{n+\lambda}|\theta) - \theta]^2 = \frac{n^2}{(n+\lambda)^2} \cdot \frac{\theta}{n} + \frac{\alpha - \lambda\theta}{n+\lambda} = \frac{n\theta + (\alpha - \lambda\theta)^2}{(n+\lambda)^2}$

$R(\theta, \bar{X}) = E(\theta - \bar{x})^2 = Var(\bar{x}|\theta) = \frac{\theta}{n}$

Compare: $\frac{\theta n + \theta^2\lambda^2 - 2\lambda\theta\alpha + \alpha^2}{(n+\lambda)^2} < \frac{\theta}{n}$

$\Rightarrow \theta n^2 + \lambda^2 n\theta^2 - 2n\lambda\theta\alpha + \alpha^2 n < \theta(n^2 + 2n\lambda + \lambda^2)$

$\Rightarrow \lambda^2 n\theta^2 - 2n\lambda\theta\alpha + \alpha^2 n < 2\lambda n\theta + \lambda^2\theta$

$\Rightarrow \underbrace{\lambda^2 n}_{a}\theta^2 \underbrace{- (2\lambda n\alpha + 2\lambda n - \lambda^2)}_{b}\theta + \underbrace{\alpha^2 n}_{c} < 0$

公式解 $= \frac{(2\lambda n\alpha + 2\lambda n - \lambda^2) \pm \sqrt{(2\lambda n\alpha + 2\lambda n - \lambda^2)^2 - 4\lambda^2\theta^2\alpha^2 n^2}}{2\lambda^2 n}$

$\Rightarrow \frac{-b - \sqrt{b^2 - 4ac}}{2a} < \theta < \frac{-b + \sqrt{b^2 - 4ac}}{2a}$ 時，$\frac{\alpha + n\bar{x}}{n+\lambda}$ 的 risk function 小於 $\bar{x}$ 的 risk function

#9

$X_1, \dots, X_n \mid \theta \overset{\text{iid}}{\sim} \exp(\theta)$ , $\theta \sim \Gamma(\alpha, \lambda)$

$\pi(\theta \mid \underline{x}) \propto \prod_{i=1}^{n} f(x_i \mid \theta)\, \pi(\theta) \propto \theta^{n} e^{-\theta \sum_{i=1}^{n} x_i}\, e^{-\lambda\theta}\, \theta^{\alpha-1} \propto \theta^{n+\alpha-1} e^{-\theta(\sum_{i=1}^{n} x_i + \lambda)}$

$\theta \mid \underline{x} \sim \Gamma(n+\alpha, \sum_{i=1}^{n} x_i + \lambda)$

∴ prior 和 posterior 皆為 Gamma, Gamma is a conjugate prior of the exponential distribution.