1 Problem 3.4.1

(a)
As given in example 3.4.7, $\Pi_2$ is spanned by $(1, \ldots, 1)'$ and $(t_1, \ldots, t_n)'$. Dimension of $\Pi_2$ is $s = 2$. So if $c_1 = (1/\sqrt{n}, \ldots, 1/\sqrt{n})'$, then by Gram-Schmidt method, a vector in $\Pi_2$ that is orthogonal to $c_1$ is given by
\[
d_2 = (t_1, \ldots, t_n)' - [(t_1, \ldots, t_n)c_1]c_1 = (t_1 - \bar{t}, \ldots, t_n - \bar{t})'
\]
\[
||d_2|| = \sum_{i=1}^{n} (t_i - \bar{t})^2 \implies c_2 = \frac{1}{\sqrt{\sum_{i=1}^{n} (t_i - \bar{t})^2}} (t_1 - \bar{t}, \ldots, t_n - \bar{t})'
\]

(b)
$s = 3$, $\Pi_2$ is spanned by $(1, \ldots, 1)'$, $(t_1, \ldots, t_n)'$ and $(t_1^2, \ldots, t_n^2)'$. With the simplifying assumptions and from (a), $c_1 = (1/\sqrt{n}, \ldots, 1/\sqrt{n})'$, $c_2 = (t_1, \ldots, t_n)'$. By Gram-Schmidt method.
\[
d_3 = (t_1^2, \ldots, t_n^2)' - [(t_1^2, \ldots, t_n^2)c_1]c_1 - [(t_1^2, \ldots, t_n^2)c_2]c_2 = (t_1^2 - t_1 \sum_{i=1}^{n} t_i^3 - \frac{1}{n}, \ldots, t_n^2 - t_n \sum_{i=1}^{n} t_i^3 - \frac{1}{n})'
\]
So, finally $c_3 = \frac{1}{||d_3||}(t_1^2 - t_1 \sum_{i=1}^{n} t_i^3 - \frac{1}{n}, \ldots, t_n^2 - t_n \sum_{i=1}^{n} t_i^3 - \frac{1}{n})'$

2 Problem 3.4.13

$\xi_\ldots = \mu, \xi_{i,} = \mu + \alpha_i, \xi_{..} = \mu + \beta_j, \xi_{..k} = \mu + \gamma_k \implies \mu = \xi_\ldots, \alpha_i = \xi_{i,} - \xi_\ldots, \beta_j = \xi_{..} - \xi_\ldots, \gamma_k = \xi_{..k} - \xi_{..}$. Note that $X_{ijk}$ is the LSE of $\xi_{ijk}$. Then by theorem 3.4.4, $\hat{\mu} = X_{\ldots}, \hat{\alpha_i} = X_{i,} - X_{\ldots}, \hat{\beta_j} = X_{..} - X_{\ldots}, \hat{\gamma_k} = X_{..k} - X_{..}$ are UMVU estimators.
Viewed as a special case of (3.4.4), $s = I + J + K - 3 = I + J + K - 2$.

3 Problem 3.4.16

(a)
In order to obtain the LSE of $\theta$, we minimize
\[
Q(\theta) = ||x - \xi||^2 = ||x - \theta A||^2 = xx^T - x(\theta A)^T - (\theta A)x^T + (\theta A)(\theta A)^T
\]
\[
\nabla Q(\theta) = -Ax^T - A\theta^T + 2AA^T \theta \theta^T = -2x^T + 2AA^T \theta \theta^T = \text{set 0}
\]
\[
\nabla Q(\theta) = \hat{\theta}AA^T \iff \hat{\theta} = xA^T (AA^T)^{-1}. \\
Thus, $\hat{\theta} = xA^T (AA^T)^{-1}$ is the LSE of $\theta$.

(b)
From Theorem 4.8, $\hat{\xi}$ is a function of $\hat{\xi}$, the least square estimate of $\xi$. From the proof of Theorem
4.4, we see that $\hat{\xi}$ is a function of the complete sufficient statistics. Hence, $\theta$ is a function of the complete sufficient statistics, so it is UMVUE for its expectation. Note that $E(\hat{\theta}) = \theta$. Thus, $\hat{\theta}$ is the UMVUE of $\theta$.

4 Problem 4.1.9

(a) Let $X_1, \ldots, X_n$ be iid Poisson distribution $P(\lambda)$, and let $\lambda$ has Gamma distribution $\Gamma(\alpha, \beta)$. Note that the expectation and variance of a RV $\lambda \sim \Gamma(\alpha, \beta)$ are $E(\lambda) = \alpha / \beta$ and $\text{Var}(\lambda) = \alpha / \beta^2$.

Moreover, also notice that $\bar{X} \sim P(\lambda)$ is the UMVUE for $\lambda$ with $E(\bar{X}) = \text{Var}(\bar{X}) = \lambda$. According to Example 4.1.3, the posterior distribution is $\pi(\lambda|\bar{x}) \sim \Gamma(\alpha + n\bar{x}, \beta + \frac{\alpha}{1 + n\beta})$. Thus, we obtain the Bayes estimator under squared error loss to be

$$
\delta_{\alpha, \beta} = E(\lambda|\bar{x}) = \frac{(g + n\bar{x})}{1 + n\alpha} \frac{\alpha}{1 + n\alpha} = \frac{g\alpha}{1 + n\alpha} + n\bar{x} \frac{\alpha}{1 + n\alpha}
$$

which is a weighted average of $g\alpha$, the estimator of $\lambda$ before any observations are taken, and $\bar{X}$, the estimator without consideration of a prior.

(b) (i) $n \to \infty$, $E(\delta_{\alpha, \beta}) \to \lambda$, $\text{Var}(\delta_{\alpha, \beta}) \to 0 \Rightarrow \delta_{\alpha, \beta} \to \lambda$ in probability.

(ii) $\alpha \to \infty$ and $\beta \to 0$, then $\delta_{\alpha, \beta} \to \bar{x}$.

Now, if both (i) and (ii) holds then $E(\delta_{\alpha, \beta}) \to \lambda$, $\text{Var}(\delta_{\alpha, \beta}) \to 0 \Rightarrow \delta_{\alpha, \beta} \to \lambda$ in probability.

5 Problem 4.1.11

$\pi(\lambda) = \frac{1}{\lambda} I(\lambda > 0)$, $f(x|\lambda) = \prod_{i=1}^{n} \left( \frac{e^{-\lambda} \lambda^x}{x_i!} \right)$. Then the posterior is $\pi(\lambda|x) \propto e^{-n\lambda} \lambda^{\sum x_i - 1}$, which is kernel of $\text{Gamma}(\sum x_i, 1/n)$. So our condition for the posterior to be proper is that $\sum_{i=1}^{n} x_i > 0$.

6 Problem 4.2.3

From (4.2.4), we know $E_\theta \delta_{\Lambda}(x) = \frac{nb^2 \theta + \sigma^2 \mu}{nb^2 + \sigma^2}$, $\text{Var}_\theta \delta_{\Lambda}(x) = \frac{nb^4 \sigma^2}{n^2 b^2 + 2nb^2 \sigma^2 + \sigma^4}$

(a) As $n \to \infty$, $E_\theta \delta_{\Lambda}(x) \to \theta$ and $\text{Var}_\theta \delta_{\Lambda}(x) \to 0$. So $\delta_{\Lambda}(x) \to \theta$ in probability as $n \to \infty$. 
(b) As $b \to 0$, $E_0\delta_\Lambda(x) \to \mu$ and $Var_0\delta_\Lambda(x) \to 0$. So $\delta_\Lambda(x) \to \mu$ in probability as $b \to 0$.

(c) For fixed $n$, as $b \to \infty$, $\delta_\Lambda(x) \to \bar{X}$ in distribution.

7 Problem 4.2.16

(a) Under squared error loss, we obtain the Bayes estimator $g(p) = p$ as

$$
\delta_\Lambda(x) = E(p|x) = \frac{\int_0^1 p^{x+a}(1-p)^{n-x+b-1}dp}{\int_0^1 p^{x+a-1}(1-p)^{n-x+b-1}dp} = \frac{x+a}{a+b+n}
$$

Under the improper prior $[p(1-p)]^{-1} I_{(0,1)}(p)$ (which implies $a = b = 0$).

To get the generalized Bayes estimator, we minimize

$$
\mathbb{E}\{(p - \delta(x))^2|x\} = \left(\frac{n}{x}\right) \int_0^1 p^{x-1}(1-p)^{n-x-1}dp
$$

For $0 < x < n$ (which makes it a proper posterior), we obtain our generalized Bayes estimator to be

$$
\frac{\int_0^1 p^x(1-p)^{n-x-1}dp}{\int_0^1 p^{x-1}(1-p)^{n-x-1}dp} = \frac{x}{n}
$$

For $x = 0$ and $x = n$, $\mathbb{E}\{(p - \delta(0))^2|x = 0\}$ and $\mathbb{E}\{(p - \delta(n))^2|x = n\}$ respectively are finite.

Thus, $\delta(x) = \frac{x}{n}$ is the generalized Bayes estimator.

(b) For $X \sim N(\theta, 1)$ and $\pi(\theta) = 1$, our posterior density becomes $\pi(\theta|x) = f(x|\theta)$. Hence,

$$
\delta(x) = E(\theta|x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \theta e^{-\frac{(\theta-x)^2}{2}}d\theta = x
$$

Thus, $\delta(x) = x$.

8 Problem 4.3.2

(a) Let $X_1, \ldots, X_n \sim \text{iid} \; \Gamma(a, b)$ where $a$ is known. Let $\eta = -\frac{1}{b}$, the density of $X$ can be expressed as

$$

f_b(x) = \exp \left\{ -\frac{1}{b} x - \log(b^a) \right\} \frac{x^{a-1}}{\Gamma(a)} I(x \geq 0) = \exp \{ \eta x - A(\eta) \} h(x)
$$
where $A(\eta) = a \cdot \log(b)$. According to (4.3.19) the conjugate prior family is

$$
\pi(\eta|k, \mu) \propto \exp \{ k \mu \eta - kA(\eta) \} = \exp \left\{ - \frac{k \mu}{b} + \log \left( \frac{1}{b \eta a} \right) \right\} = \frac{1}{b \eta a} e^{- \frac{k \mu}{b}}
$$

which we recognize as the kernel of $IG(ka - 1, k\mu)$ (for $ka > 1$ and $k\mu > 0$ under the support $b > 0$). Thus, the conjugate prior for $\eta = -\frac{1}{b}$ is equivalent to an inverted gamma on $b$.

**Note:** Let $b \sim \text{InvGam}(\alpha, \beta) \Rightarrow p_{\alpha,\beta}(b) = \frac{\beta^\alpha}{\Gamma(\alpha)} \left( \frac{1}{b} \right)^{\alpha+1} e^{-\frac{\beta}{b}} I(b > 0)$ where $E(b) = \frac{\beta}{\alpha-1}$.

Moreover,

$$
E \left( \frac{1}{b^2} \right) = \frac{\beta^\alpha}{\Gamma(\alpha)} \int^\infty_{0} \frac{1}{b^{\alpha+2}} e^{-\frac{\beta}{b}} db = \frac{\beta^{\alpha}}{\Gamma(\alpha)} \frac{\Gamma(\alpha+2)}{\beta^{\alpha+2}} = \frac{\alpha}{\beta}
$$

(which makes senses since $\frac{1}{b} \sim \text{Gamma}(\alpha, \frac{1}{\beta}) \Rightarrow E(\frac{1}{b}) = \frac{\alpha}{\beta}$) and

$$
E \left( \frac{1}{b^2} \right) = \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+2)}{\beta^{\alpha+2}} = \frac{\alpha}{\beta}
$$

Let $X_1, \ldots, X_n \sim \text{iid } \Gamma(a, b)$ where $b \sim \text{InvGam}(c, d)$. Note that $\pi(b|x) \propto \frac{1}{b^{an+c+1}} e^{-\frac{1}{b}(\sum x_i + d)}$ which implies $b|x \sim \text{InvGam}(an + c, \sum x_i + d)$. First, let’s consider the case where $L(b, \delta) = (b - \delta)^2$. From Corollary 4.1.2

$$
\delta(x) = E(b|x) = \frac{\sum x_i + d}{an + c - 1}
$$

Lastly, let’s consider the case where $L(b, \delta) = \frac{1}{b^2}(b - \delta)^2$. Again, from Corollary 4.1.2

$$
\delta(x) = \frac{E(\frac{1}{b^2}|x)}{E(\frac{1}{b^2})} = \frac{\sum x_i + d}{an + c + 1}
$$

**9 Problem 4.3.8**

**(a)**

Note: from Lemma 1.5.15, Stein’s indentity is given to be

$$
E \left\{ \left[ \frac{h'(X)}{h(X)} + \sum_{i=1}^{s} \eta_i T'_i(X) \right] g(X) \right\} = - Eg'(X)
$$

Let $g(X) = 1 \Rightarrow g'(X) = 0$. From Stein’s indentity,

$$
E_n \left\{ \left[ \frac{h'(X)}{h(X)} + \sum_{i=1}^{s} \eta_i T'_i(X) \right] g(X) \right\} = 0
$$
(where \( \frac{h'(X)}{h(X)} = \frac{\delta}{\delta x_j} \log(h(X)) \) and \( T'_i = \frac{\delta}{\delta x_j} T_i(X) \))

\[
\Rightarrow \sum_{i=1}^{s} \eta_i E_{\eta}(\frac{\delta}{\delta x_j} T_i(X)) = E_{\eta}(\frac{\delta}{\delta x_j} \log(h(X)))
\]

(b) Let \( X_i \sim iid \Gamma(a, b) \) where \( a \) is known. The density is

\[
f(x|b) = \frac{\text{Gamma}(a, b)}{(\Gamma(a))^a} \exp \left\{ -a \cdot \log(b) - \frac{1}{b} \sum x_i \right\} = h(x) \cdot \exp \left\{ -A(\eta) + \eta \sum T_i \right\}
\]

that is, \( \eta_i = \frac{1}{b} \) and \( T_i = -x_i \). Notice that

\[
E_{\eta}(\frac{\delta}{\delta x_j} \log(h(X))) = \sum_{i=1}^{s} \eta_i E_{\eta}(\frac{\delta}{\delta x_j} T_i(X)) = \frac{1}{b} \sum E(\frac{\delta}{\delta x_j}(-x_i)) = -\frac{1}{b}
\]

which implies \( \frac{\delta}{\delta x_j} \log(h(X)) \) is an unbiased estimator of \( \frac{1}{b} \). Let’s consider

\[
\frac{\delta}{\delta x_j} \log(h(X)) = \frac{\delta}{\delta x_j} \left( -n \log(\Gamma(a)) + (a - 1) \sum \log(x_i) \right) = \frac{a - 1}{x_j}
\]

Thus, \( \frac{a - 1}{x_j} \) is an unbiased estimator of \( \frac{1}{b} \) for all \( j \). Moreover, \( (a - 1) \sum (x_j^{-1} / n) \) is also unbiased for \( 1/b \).

(c) Let \( X_i \sim iid \text{Beta}(a, b) \). Let’s assume \( b \) is known. Using part(a), we have

\[
(a - 1)E(\frac{1}{x_j}) = (b - 1)E(\frac{1}{1-x_j})
\]

Note that \( E(\frac{1}{x_j}) = \frac{a + b - 1}{a - 1} \Rightarrow (a + b - 1) = (b - 1)E(\frac{1}{1-x_j}) \Rightarrow a = \frac{(b - 1)}{n} E(\sum \frac{x_j}{1-x_j})
\]

Thus, the unbiased estimator for \( a \) is \( \frac{(b - 1)}{n} (\sum \frac{x_j}{1-x_j}) \).

Similar argument is applied to the case where \( a \) is known.

10 Problem 4.3.12

(a)

\[
E[(a \bar{X} + b) - \mu]^2 = a^2 E(\bar{X} - \mu)^2 + E[b + (a - 1)\mu]^2 + 2aE(\bar{X} - \mu)[b + (a - 1)\mu]
\]

\[
= a^2 Var(\bar{X}) + [(a - 1)\mu + b]^2
\]

(b) If \( \mu \) is unbounded, then \( MSE(a \bar{X} + b) = a^2 Var(\bar{X}) + [(a - 1)\mu + b]^2 \) is also unbounded if \( a \neq 1 \).

(c) If \( \mu \) is bounded, the Bayes estimator can have finite MSE.
11 Problem 4.4.1

(a) $X \mid p \sim \text{Bin}(n, p)$, $p \sim \text{Beta}(\alpha, \alpha)$. Then $p \mid X \sim \text{Beta}(\alpha + x, \alpha + n - x)$. We know $g(X) = n - X$, $\bar{g}(p) = 1 - p$, $g^*(u) = 1 - u$ and $\delta(X) = E(p \mid X) = \frac{\alpha + x}{2\alpha + n}$. So $g^*(\delta(X)) = 1 - \frac{\alpha + x}{2\alpha + n} = \frac{\alpha + n - x}{2\alpha + n} = \delta(g(X))$. So, this Bayes rule is equivariant.

(b) 

$$g^*(\delta(X)) = 1 - \delta(X) = \frac{\int_0^1 p^x(1 - p)^{n - x} \pi(p) dp - \int_0^1 p^{x+1}(1 - p)^{n - x} \pi(p) dp}{\int_0^1 p^x(1 - p)^{n - x} \pi(p) dp}$$

Let $p_1 = 1 - p$ then

$$g^*(\delta(X)) = \frac{\int_0^1 (p_1)^{n - x + 1}(1 - p_1)^x \pi(1 - p_1) dp_1}{\int_0^1 (p_1)^{n - x}(1 - p_1)^x \pi(1 - p_1) dp_1}$$

But $\pi(\cdot)$ is symmetric about $1/2$, so $\pi(1 - p_1) = \pi(p_1)$. Then

$$g^*(\delta(X)) = \frac{\int_0^1 (p_1)^{n - x + 1}(1 - p_1)^x \pi(p_1) dp_1}{\int_0^1 (p_1)^{n - x}(1 - p_1)^x \pi(p_1) dp_1} = \delta(g(X))$$