Statistical Inference: Complete Solutions

Contents

1 Problem 1: Exponential Distribution and EM Algorithm

A factory produces n bulbs using a machine. Assume the observed data $X_1, \dots, X_n$ record the failure times of the bulbs and $X_i \sim \text{Exp}(\lambda)$ with constant failure rate λ. Note the pdf for X_i is $f(x; \lambda) = \lambda e^{-\lambda x}$ with mean $1/\lambda$ and variance $(1/\lambda)^2$.

1.1 Part (a): Asymptotic Inference on β^2

Consider $\beta = 1/\lambda$. We would like to make asymptotic inference on the variance β^2.

1.1.1 Part (i): MLE and Asymptotic Distribution

Deriving the MLE:

Since $\beta = 1/\lambda$, we have $\lambda = 1/\beta$. The pdf in terms of β is:

$$f(x; \beta) = \frac{1}{\beta} e^{-x/\beta}, \quad x > 0$$

The likelihood function is:

$$L(\beta) = \prod_{i=1}^{n} \frac{1}{\beta} e^{-X_i/\beta} = \beta^{-n} \exp\left(-\frac{1}{\beta} \sum_{i=1}^{n} X_i\right)$$

The log-likelihood is:

$$\ell(\beta) = -n \ln \beta - \frac{1}{\beta} \sum_{i=1}^{n} X_i$$

Taking the derivative and setting it to zero:

$$\frac{d\ell}{d\beta} = -\frac{n}{\beta} + \frac{1}{\beta^2} \sum_{i=1}^{n} X_i = 0$$

Solving:

$$\boxed{\hat{\beta}_n = \bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i}$$

Asymptotic Distribution:
Theorem Used: Central Limit Theorem (CLT)
Conditions:

1. $X_1, \dots, X_n$ are i.i.d. ✓
2. $\mathbb{E}[X_i] = \beta$ (finite mean) ✓
3. $\text{Var}(X_i) = \beta^2$ (finite variance) ✓

Application: Since $\hat{\beta}_n = \bar{X}_n$, by the CLT:

$$\boxed{\sqrt{n}(\hat{\beta}_n - \beta) \xrightarrow{d} N(0, \beta^2)}$$

1.1.2 Part (ii): Asymptotic Normality of $\hat{\beta}_n^2$

Theorem Used: Delta Method

Statement: If $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, \sigma^2)$ and g is continuously differentiable with $g'(\theta) \neq 0$, then:

$$\sqrt{n}(g(\hat{\theta}_n) - g(\theta)) \xrightarrow{d} N(0, [g'(\theta)]^2\sigma^2)$$

Conditions:

1. $\sqrt{n}(\hat{\beta}_n - \beta) \xrightarrow{d} N(0, \beta^2)$ ✓(from part i)
2. Let $g(\beta) = \beta^2$, which is continuously differentiable ✓
3. $g'(\beta) = 2\beta \neq 0$ for $\beta > 0$ ✓

Application:
With $g(\beta) = \beta^2$ and $g'(\beta) = 2\beta$:
The asymptotic variance is:

$$[g'(\beta)]^2 \cdot \beta^2 = (2\beta)^2 \cdot \beta^2 = 4\beta^4$$

Therefore:

$$\boxed{\sqrt{n}(\hat{\beta}_n^2 - \beta^2) \xrightarrow{d} N(0, 4\beta^4)}$$

1.2 Part (b): EM Algorithm for Mixture of Exponentials

Now we find that the factory has used two machines: Machine A and Machine B. For each bulb, the machine used is chosen at random: with probability π, the bulb is made by Machine A, and with probability $1 - \pi$, it is made by Machine B. Machine A produces bulbs with failure rate λ_A and Machine B produces bulbs with failure rate λ_B.

The latent variables $z_1, \ldots, z_n$ are not observed, where $z_i = 1$ if the i-th bulb is from Machine A and $z_i = 0$ if from Machine B.

1.2.1 Part (i): Complete Data Log-Likelihood

The complete data likelihood for observation i is:

$$L_i = [\pi\lambda_A e^{-\lambda_A x_i}]^{z_i}[(1 - \pi)\lambda_B e^{-\lambda_B x_i}]^{1-z_i}$$

The complete data log-likelihood is:

$$\boxed{\ell_C(\theta \mid X, Z) = \sum_{i=1}^{n} \Big[z_i \ln \pi + (1 - z_i)\ln(1 - \pi) + z_i \ln \lambda_A - z_i\lambda_A x_i + (1 - z_i)\ln \lambda_B - (1 - z_i)\lambda_B x_i\Big]}$$

1.2.2 Part (ii): Calculate $\gamma_i^{(k)}$ and $\mathrm{Pr}(Z \mid X, \theta^{(k)})$

Using **Bayes' Rule**:

$$\gamma_i^{(k)} = \Pr(z_i = 1 \mid x_i, \theta^{(k)}) = \frac{\Pr(x_i \mid z_i = 1, \theta^{(k)})\Pr(z_i = 1 \mid \theta^{(k)})}{\Pr(x_i \mid \theta^{(k)})}$$

$$\boxed{\gamma_i^{(k)} = \frac{\pi^{(k)}\lambda_A^{(k)}e^{-\lambda_A^{(k)}x_i}}{\pi^{(k)}\lambda_A^{(k)}e^{-\lambda_A^{(k)}x_i} + (1-\pi^{(k)})\lambda_B^{(k)}e^{-\lambda_B^{(k)}x_i}}}$$

Since $(x_1, z_1), \ldots, (x_n, z_n)$ are mutually independent:

$$\boxed{\Pr(Z \mid X, \theta^{(k)}) = \prod_{i=1}^{n}(\gamma_i^{(k)})^{z_i}(1-\gamma_i^{(k)})^{1-z_i}}$$

1.2.3 Part (iii): E-Step

Taking the expectation of the complete data log-likelihood with respect to $Z \mid X, \theta^{(k)}$:

Since $\mathbb{E}[z_i \mid x_i, \theta^{(k)}] = \gamma_i^{(k)}$, we substitute $z_i \to \gamma_i^{(k)}$:

$$\boxed{Q(\theta, \theta^{(k)}, X) = \sum_{i=1}^{n}\left[\gamma_i^{(k)}\ln\pi + (1-\gamma_i^{(k)})\ln(1-\pi) + \gamma_i^{(k)}(\ln\lambda_A - \lambda_A x_i) + (1-\gamma_i^{(k)})(\ln\lambda_B - \lambda_B x_i)\right]}$$

1.2.4 Part (iv): M-Step Update Equations

Maximizing Q with respect to each parameter:

For π:

$$\boxed{\pi^{(k+1)} = \frac{1}{n}\sum_{i=1}^{n}\gamma_i^{(k)}}$$

For λ_A:

$$\boxed{\lambda_A^{(k+1)} = \frac{\sum_{i=1}^{n}\gamma_i^{(k)}}{\sum_{i=1}^{n}\gamma_i^{(k)}x_i}}$$

For λ_B:

$$\boxed{\lambda_B^{(k+1)} = \frac{\sum_{i=1}^{n}(1-\gamma_i^{(k)})}{\sum_{i=1}^{n}(1-\gamma_i^{(k)})x_i}}$$

2 Problem 2: Likelihood Ratio and Score Tests for Poisson Model

Consider random variables Y_i ($i = 1, \dots, n$), the number of hospitalizations of patient i in the past five years, and X_i ($i = 1, \dots, n$) is the treatment covariate of patient i, where $X_i = 0$ means patient i is a control and $X_i = 1$ means patient i is treated. Assume that $Y_1, \dots, Y_n$ are independent. For $i = 1, \dots, n$, the random variable Y_i follows Poisson(μ_i).

Use notation: $C = \{i : X_i = 0\}$ and $D = \{i : X_i = 1\}$ with sizes n_0 and n_1, respectively. Also denote $\bar{Y}_C = (\sum_{i \in C} Y_i)/n_0$, $\bar{Y}_D = (\sum_{i \in D} Y_i)/n_1$, and $\bar{Y} = (\sum_{i=1}^n Y_i)/n$.

2.1 Part (a): Likelihood Ratio Test under Setup (A)

Setup (A): $\mu_i = \eta_0$ if $X_i = 0$ and $\mu_i = \eta_1$ if $X_i = 1$.

Test: $H_0^A : \eta_0 = \eta_1$ vs. $H_1^A : \eta_0 \neq \eta_1$

2.1.1 Step 1: Write the Likelihood

The full likelihood under the alternative (unrestricted) model is:

$$L(\eta_0, \eta_1) = \prod_{i \in C} \frac{\eta_0^{Y_i} e^{-\eta_0}}{Y_i!} \cdot \prod_{i \in D} \frac{\eta_1^{Y_i} e^{-\eta_1}}{Y_i!}$$

The log-likelihood is:

$$\ell(\eta_0, \eta_1) = \sum_{i \in C} (Y_i \ln \eta_0 - \eta_0) + \sum_{i \in D} (Y_i \ln \eta_1 - \eta_1) + \text{const}$$

2.1.2 Step 2: MLEs under H_1^A

$$\hat{\eta}_0 = \bar{Y}_C, \qquad \hat{\eta}_1 = \bar{Y}_D$$

2.1.3 Step 3: MLE under H_0^A

Under H_0 ($\eta_0 = \eta_1 = \eta$):

$$\tilde{\eta} = \bar{Y}$$

2.1.4 Step 4: Likelihood Ratio Statistic

$$\boxed{T_{LR} = 2 \left[n_0 \bar{Y}_C \ln \frac{\bar{Y}_C}{\bar{Y}} + n_1 \bar{Y}_D \ln \frac{\bar{Y}_D}{\bar{Y}} \right]}$$

2.1.5 Step 5: Asymptotic Distribution

By **Wilks' Theorem**, under H_0^A:

$$\boxed{T_{LR} \xrightarrow{d} \chi_1^2}$$

2.2 Part (b): Score Test under Setup (B)

Setup (B): $\mu_i = a + bX_i$

Test: $H_0^B : b = 0$ vs. $H_1^B : b \neq 0$

Parameters: $\theta = (b, a)$ with b as parameter of interest, a as nuisance.

2.2.1 Part (i): Constrained MLE under H_0^B

Under H_0^B (where $b = 0$), we have $\mu_i = a$ for all observations.

$$\boxed{\tilde{\theta} = (\tilde{b}, \tilde{a}) = (0, \bar{Y})}$$

2.2.2 Part (ii): Total Information Matrix

The log-likelihood for the full model ($\mu_i = a + bX_i$):

$$\ell(b, a) = \sum_{i=1}^{n} [Y_i \ln(a + bX_i) - (a + bX_i)] + \text{const}$$

Expected Fisher Information:

$$I_{11} = \sum_{i=1}^{n} \frac{X_i^2}{a + bX_i} = \frac{n_1}{a + b}$$
$$I_{22} = \sum_{i=1}^{n} \frac{1}{a + bX_i} = \frac{n_0}{a} + \frac{n_1}{a + b}$$
$$I_{12} = I_{21} = \sum_{i=1}^{n} \frac{X_i}{a + bX_i} = \frac{n_1}{a + b}$$

$$\boxed{I_T(\theta) = \begin{pmatrix} \frac{n_1}{a+b} & \frac{n_1}{a+b} \\ \frac{n_1}{a+b} & \frac{n_0}{a} + \frac{n_1}{a+b} \end{pmatrix}}$$

2.2.3 Part (iii): Score Test Statistic

Score for b at $\tilde{\theta}$:

$$S_b(\tilde{\theta}) = \frac{n_1(\bar{Y}_D - \bar{Y})}{\bar{Y}}$$

Using the hint $I^{11} = (I_{11} - I_{12}I_{22}^{-1}I_{21})^{-1}$:

$$I^{11} = \frac{n\bar{Y}}{n_0 n_1}$$

Score test statistic:

$$\boxed{T_S = \frac{n \cdot n_1(\bar{Y}_D - \bar{Y})^2}{n_0 \bar{Y}}}$$

Asymptotic Distribution under H_0^B:

$$\boxed{T_S \xrightarrow{d} \chi_1^2}$$

3 Problem 3: Joint Density, MLE, and Influence Function

Consider a pair of random variables (X, Y) where X is discrete with $\Pr[X = 1] = \Pr[X = 4] = \frac{1}{2}$ and $[Y|X = x] \sim N(\theta, x)$.

3.1 Part (i): Joint Density and Identifiability

3.1.1 Deriving the Joint Density

The joint density/pmf is $f(x, y; \theta) = f(y|x; \theta) \cdot P(X = x)$

Since $Y|X = x \sim N(\theta, x)$:

$$f(y|x;\theta) = \frac{1}{\sqrt{2\pi x}} \exp\left\{-\frac{(y-\theta)^2}{2x}\right\}$$

With $P(X = x) = \frac{1}{2}$ for $x \in \{1, 4\}$:

$$\boxed{f(x,y;\theta) = \frac{1}{2\sqrt{2\pi x}} \exp\left\{-\frac{(y-\theta)^2}{2x}\right\}}$$

3.1.2 Verifying Identifiability

Identifiability condition: $\theta_1 \neq \theta_2 \Rightarrow f(x, y; \theta_1) \neq f(x, y; \theta_2)$ for some (x, y).

Suppose $f(x, y; \theta_1) = f(x, y; \theta_2)$ for all (x, y). Then $(y-\theta_1)^2 = (y-\theta_2)^2$ for all y, which requires $\theta_1 = \theta_2$.

$$\boxed{\text{Therefore, } \theta \text{ is identifiable.}}$$

3.2 Part (ii): MLE and Limiting Distribution

3.2.1 Finding the MLE

The log-likelihood for sample $\{(x_i, y_i)\}_{i=1}^n$:

$$\ell(\theta) = \sum_{i=1}^n \left[-\frac{1}{2}\ln(2\pi x_i) - \frac{(y_i-\theta)^2}{2x_i}\right] + \text{const}$$

Taking the derivative and setting to zero:

$$\boxed{\hat{\theta}_n = \frac{\sum_{i=1}^n y_i/x_i}{\sum_{i=1}^n 1/x_i}}$$

This is a **weighted least squares** estimator with weights $w_i = 1/x_i$.

3.2.2 Limiting Distribution

Fisher Information:

$$I_1(\theta) = \mathbb{E}\left[\frac{1}{X}\right] = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{1}{4} = \frac{5}{8}$$

By asymptotic normality of MLEs:

$$\boxed{\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N\left(0, \frac{8}{5}\right)}$$

3.3 Part (iii): Influence Function

The score function is $\psi(x, y; \theta) = \frac{y-\theta}{x}$
We have:

- $\frac{\partial \psi}{\partial \theta} = -\frac{1}{x}$
- $\mathbb{E}\left[\frac{\partial \psi}{\partial \theta}\right] = -\mathbb{E}\left[\frac{1}{X}\right] = -\frac{5}{8}$

Therefore:

$$\boxed{IF(x, y; \theta) = \frac{8(y-\theta)}{5x}}$$

4 Problem 4: Beta Distribution and Sample Median

Let $X_1, \ldots, X_n$ be a sample from the beta distribution with density $f(x;\theta) = \theta x^{\theta-1}$, where $0 < x < 1$ and $\theta > 0$.

4.1 Part (i): Population Median and Limiting Distribution

4.1.1 Finding the Population Median

The CDF is $F(x) = x^\theta$. The median η satisfies $F(\eta) = 0.5$:

$$\eta^\theta = 0.5 \implies \boxed{\eta = m(\theta) = 2^{-1/\theta}}$$

4.1.2 Limiting Distribution of $\sqrt{n}(\hat{\eta}_n - \eta)$

The density at the median:

$$f(\eta) = \theta\eta^{\theta-1} = \theta \cdot 2^{1/\theta-1}$$

The asymptotic variance:

$$\sigma_\eta^2 = \frac{1}{4[f(\eta)]^2} = \frac{1}{\theta^2 \cdot 2^{2/\theta}}$$

$$\boxed{\sqrt{n}(\hat{\eta}_n - \eta) \xrightarrow{d} N\left(0, \frac{1}{\theta^2 \cdot 2^{2/\theta}}\right)}$$

4.2 Part (ii): Consistency of $\hat{\theta}_n$

Since $\eta = 2^{-1/\theta}$, we have $\theta = \frac{\log(1/2)}{\log \eta}$.

Since $\hat{\eta}_n \xrightarrow{p} \eta$, by the **Continuous Mapping Theorem**:

$$\boxed{\hat{\theta}_n = \frac{\log(1/2)}{\log(\hat{\eta}_n)} \xrightarrow{p} \frac{\log(1/2)}{\log \eta} = \theta}$$

4.3 Part (iii): Influence Function and Limiting Distribution of $\hat{\theta}_n$

Let $g(\eta) = \frac{\log(1/2)}{\log \eta}$, so $\hat{\theta}_n = g(\hat{\eta}_n)$.

Computing $g'(\eta)$:

$$g'(\eta) = \frac{\theta^2}{2^{-1/\theta}\log 2}$$

Influence Function:

$$\boxed{IF_{\hat{\theta}_n}(x) = \frac{\theta(1 - 2 \cdot I\{x \leq \eta\})}{\log 2}}$$

Limiting Distribution:

$$\boxed{\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N\left(0, \frac{\theta^2}{(\log 2)^2}\right)}$$

4.4 Part (iv): One-Step Newton-Raphson Estimator

Log-likelihood: $\ell(\theta) = n\log\theta + (\theta - 1)\sum_{i=1}^{n}\log X_i$
Score: $\ell'(\theta) = \frac{n}{\theta} + \sum_{i=1}^{n}\log X_i$
Observed Information: $-\ell''(\theta) = \frac{n}{\theta^2}$
The one-step estimator:

$$\boxed{\tilde{\theta}_n = 2\hat{\theta}_n + \frac{\hat{\theta}_n^2}{n}\sum_{i=1}^{n}\log X_i}$$

4.5 Part (v): Limiting Distribution of $\tilde{\theta}_n$

Since $\tilde{\theta}_n$ is asymptotically efficient, its asymptotic variance equals the inverse Fisher information.
Fisher Information: $I(\theta) = \frac{1}{\theta^2}$

$$\boxed{\sqrt{n}(\tilde{\theta}_n - \theta) \xrightarrow{d} N(0, \theta^2)}$$

5 Problem 5: Generalized Linear Model and Quasi-Likelihood

Consider the following generalized linear model. For $i = 1, \ldots, n$:

- $\mathbb{E}(y_i) = \mu_i$
- $\log(\mu_i) = \alpha + x_i\beta$ (log link), so $\mu_i = e^{\alpha + x_i\beta}$
- $\text{Var}(y_i) = \sigma^2\mu_i$
- $\beta \in \mathbb{R}$ (scalar)

5.1 Part (i): Maximum Quasi-Likelihood Estimating Equations

For a GLM with variance function $V(\mu_i) = \sigma^2\mu_i$, the quasi-score is:

$$U(\theta) = \sum_{i=1}^{n} \frac{(y_i - \mu_i)}{\text{Var}(y_i)} \cdot \frac{\partial \mu_i}{\partial \theta}$$

Since $\mu_i = e^{\alpha + x_i\beta}$:

$$\frac{\partial \mu_i}{\partial \alpha} = \mu_i, \qquad \frac{\partial \mu_i}{\partial \beta} = x_i\mu_i$$

Quasi-Score Equations:
For α:

$$\boxed{\sum_{i=1}^{n}(y_i - e^{\hat{\alpha} + x_i\hat{\beta}}) = 0}$$

For β:

$$\boxed{\sum_{i=1}^{n} x_i(y_i - e^{\hat{\alpha} + x_i\hat{\beta}}) = 0}$$

In matrix form with $\mathbf{X} = (\mathbf{1}, \mathbf{x})$:

$$\boxed{\mathbf{X}^T(\mathbf{y} - \boldsymbol{\mu}) = \mathbf{0}}$$

5.2 Part (ii): Estimator for σ^2

Using the Pearson chi-squared statistic:

$$\boxed{\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^{n} \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i}}$$

where $\hat{\mu}_i = e^{\hat{\alpha} + x_i\hat{\beta}}$.

5.3 Part (iii): Asymptotic Distribution of $\hat{\theta}$

Sensitivity Matrix:

$$\mathbf{S}(\theta) = \frac{1}{\sigma^2} \sum_{i=1}^{n} \mu_i \begin{pmatrix} 1 & x_i \\ x_i & x_i^2 \end{pmatrix}$$

Since $\mathbf{V} = \mathbf{S}$ for this model:

$$\boxed{\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N\left(\mathbf{0}, n\mathbf{S}(\theta)^{-1}\right)}$$

Or equivalently:

$$\boxed{\hat{\theta} \overset{\text{approx}}{\sim} N\left(\theta, \sigma^2 \left(\sum_{i=1}^{n} \mu_i \begin{pmatrix} 1 & x_i \\ x_i & x_i^2 \end{pmatrix}\right)^{-1}\right)}$$

6 Problem 6: Random Effects Model

Consider the model:

$$y_{ij} = \alpha + u_i + \epsilon_{ij}$$

where $i = 1, \ldots, m$, $j = 1, \ldots, n_i$; $u_i \overset{\text{iid}}{\sim} N(0, \sigma_u^2)$; $\epsilon_{ij} \overset{\text{iid}}{\sim} N(0, \sigma^2)$; u_i's and ϵ_{ij}'s are independent.

6.1 Part (i): MLE for α when $\sigma_u^2 = \rho\sigma^2$

Let $\mathbf{y} = (\mathbf{y}_1^T, \ldots, \mathbf{y}_m^T)^T$ with total $N = \sum_{i=1}^m n_i$ observations.

Mean: $\mathbb{E}[\mathbf{y}] = \alpha \mathbf{1}_N$

Covariance:

$$\mathbf{V} = \sigma^2 \cdot \text{diag}(\rho \mathbf{J}_{n_1} + \mathbf{I}_{n_1}, \ldots, \rho \mathbf{J}_{n_m} + \mathbf{I}_{n_m})$$

where $\mathbf{J}_{n_i} = \mathbf{1}_{n_i} \mathbf{1}_{n_i}^T$.

MLE for α:

$$\boxed{\hat{\alpha}_{MLE} = (\mathbf{1}_N^T \mathbf{V}^{-1} \mathbf{1}_N)^{-1} \mathbf{1}_N^T \mathbf{V}^{-1} \mathbf{y}}$$

Justification: This is the GLS estimator, which equals the MLE for multivariate normal data.

6.2 Part (ii): Unbiasedness of REML Estimator $\hat{\sigma}^2$

The REML estimator is:

$$\hat{\sigma}^2 = \frac{1}{\sum_{i=1}^m (n_i - 1)} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i\cdot})^2$$

Key Observation:

$$y_{ij} - \bar{y}_{i\cdot} = \epsilon_{ij} - \bar{\epsilon}_{i\cdot}$$

The random effect u_i cancels out!

Since $\epsilon_{ij} \overset{\text{iid}}{\sim} N(0, \sigma^2)$:

$$\mathbb{E}\left[\sum_{j=1}^{n_i} (\epsilon_{ij} - \bar{\epsilon}_{i\cdot})^2\right] = (n_i - 1)\sigma^2$$

Therefore:

$$\boxed{E[\hat{\sigma}^2] = \frac{\sum_{i=1}^m (n_i - 1)\sigma^2}{\sum_{i=1}^m (n_i - 1)} = \sigma^2}$$

6.3 Part (iii): Best Predictor $BP(u_i)$

The joint distribution of $(u_i, \mathbf{y}_i)$ is:

$$\begin{pmatrix} u_i \\ \mathbf{y}_i \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ \alpha \mathbf{1}_{n_i} \end{pmatrix}, \begin{pmatrix} \sigma_u^2 & \sigma_u^2 \mathbf{1}_{n_i}^T \\ \sigma_u^2 \mathbf{1}_{n_i} & \sigma_u^2 \mathbf{J}_{n_i} + \sigma^2 \mathbf{I}_{n_i} \end{pmatrix} \right)$$

Using the conditional distribution formula:

$$\boxed{BP(u_i) = E[u_i \mid \mathbf{y}_i] = \frac{n_i \sigma_u^2}{n_i \sigma_u^2 + \sigma^2} (\bar{y}_{i\cdot} - \alpha)}$$

Interpretation: This is a shrinkage estimator that pulls $(\bar{y}_{i\cdot} - \alpha)$ toward zero.

7 Problem 7: ANOVA and Model Comparison

Researchers investigate the effect of exercise treatment on systolic blood pressure (SBP). Patients are randomized into 4 groups with varying exercise durations: 0, 30, 60, and 90 minutes per day. For each patient j in group i ($i = 0, 1, 2, 3$; $j = 1, \ldots, 5$), they also measure their baseline BMI, denoted x_{ij}. Define d_i as the numerical exercise duration level.

Models:

Model	Description	Model Equation	SSE
A	Categorical treatment + BMI	$y_{ij} = \mu + \alpha_i + \theta x_{ij} + \epsilon_{ij}$	SSE_A
B	Categorical treatment only	$y_{ij} = \mu + \alpha_i + \epsilon_{ij}$	SSE_B
C	Linear treatment + BMI	$y_{ij} = \mu + \delta d_i + \theta x_{ij} + \epsilon_{ij}$	SSE_C
D	Linear treatment only	$y_{ij} = \mu + \delta d_i + \epsilon_{ij}$	SSE_D
E	Mean model	$y_{ij} = \mu + \epsilon_{ij}$	SSE_E

7.1 Part (a): Model D is Nested within Model B

Model B: $y_{ij} = \mu + \alpha_i + \epsilon_{ij}$ (categorical treatment only)

Model D: $y_{ij} = \mu + \delta d_i + \epsilon_{ij}$ (linear treatment only)

Constraint to obtain Model D from Model B: Set $\alpha_i = \delta d_i$ for all i.

With $d_0 = 0, d_1 = 30, d_2 = 60, d_3 = 90$:

- $\alpha_0 = 0$
- $\alpha_1 = 30\delta$
- $\alpha_2 = 60\delta$
- $\alpha_3 = 90\delta$

$$\boxed{\text{Model D is nested in Model B because setting } \alpha_i = \delta d_i \text{ reduces Model B to Model D.}}$$

7.2 Part (b): Sequential Sums of Squares

7.2.1 (i) $R(\mu)$

$$\boxed{R(\mu) = \text{TSS} - \text{SSE}_E}$$

7.2.2 (ii) $R(\boldsymbol{\alpha} \mid \mu, \delta)$

$$\boxed{R(\boldsymbol{\alpha} \mid \mu, \delta) = \text{SSE}_D - \text{SSE}_B}$$

7.2.3 (iii) $R(\theta \mid \boldsymbol{\alpha})$

$$\boxed{R(\theta \mid \boldsymbol{\alpha}) = \text{SSE}_B - \text{SSE}_A}$$

7.2.4 (iv) $R(\boldsymbol{\alpha} \mid \mu, \delta, \theta)$

$$\boxed{R(\boldsymbol{\alpha} \mid \mu, \delta, \theta) = \text{SSE}_C - \text{SSE}_A}$$

7.3 Part (c): Hypothesis Tests at $\alpha = 0.05$

7.3.1 Test 1: Does BMI explain significant variability in SBP?

Compare Model B vs Model A

Hypotheses:

$$H_0 : \theta = 0 \quad \text{vs.} \quad H_1 : \theta \neq 0$$

Test Statistic:

$$\boxed{F = \frac{\text{SSE}_B - \text{SSE}_A}{\text{SSE}_A/15} \sim F_{1,15} \text{ under } H_0}$$

where $df_1 = 1$ and $df_2 = 20 - 5 = 15$.
Decision: Reject H_0 if $F > F_{1,15,0.05} \approx 4.54$

7.3.2 Test 2: Is the linear model adequate?

Compare Model A vs Model C

Hypotheses:

$$H_0 : \text{Linear model is adequate} \quad \text{vs.} \quad H_1 : \text{Linear model has lack of fit}$$

Test Statistic:

$$\boxed{F = \frac{(\text{SSE}_C - \text{SSE}_A)/2}{\text{SSE}_A/15} \sim F_{2,15} \text{ under } H_0}$$

where $df_1 = 5 - 3 = 2$ and $df_2 = 15$.
Decision: Reject H_0 (conclude linear model is inadequate) if $F > F_{2,15,0.05} \approx 3.68$