STA260 Summer 2024 Tutorial 11 (11.1, 11.2, 11.3, 11.4)

Disclaimer: STA260 covers extensive material. Tutorials serve as additional aids, but they cannot cover every question type.

Relevant Review from Lecture

Simple Linear Regression model: $y = \beta_0 + \beta_1 x + \epsilon$, $\epsilon \sim N(0, \sigma^2)$
Sum of Squares (To be Minimized): $SSE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$
Some Relevant Formulas:

1. $\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$ 2. $\hat{\beta}_0 = \overline{y} - \hat{\beta}_1\overline{x}$ 3. $V(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}$ 4. $V(\hat{\beta}_0) = \sigma^2\left(\frac{1}{n} + \frac{\overline{x}^2}{S_{xx}}\right)$
5. $S_{xy} = \sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) = \sum_{i=1}^{n} x_i y_i - n(\overline{x})(\overline{y})$
6. $S_{xx} = \sum_{i=1}^{n}(x_i - \overline{x})^2 = \sum_{i=1}^{n} x_i^2 - n\overline{x}^2$

Question 1 (11.1 from the textbook)

If $\hat{\beta}_0$ and $\hat{\beta}_1$ are the least-squares estimates for the intercept and the slope in a simple linear regression model, show that the least-squares equation $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ always goes through the point $(\bar{x}, \bar{y})$.
Hint: substitute $\bar{x}$ for x in the least-squares equation and use the fact that $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$

Question 2 (11.10 from the textbook)

Suppose we have the postulated the model:

$$Y_i = \beta_1 x_i + \epsilon_i, \quad i = 1, 2, ..., n,$$

where the ϵ_i's are independent and identically distributed random variables with $\mathbb{E}(\epsilon_i) = 0$. Then $\hat{y}_i = \hat{\beta}_1 x_i$ is the predicted value of y when $x = x_i$ and $SSE = \sum_{i=1}^{n} [y_i - \hat{\beta}_1 x_i]^2$. Find the least-squares estimator of β_1. (Notice that the equation $y = \beta x$ describes a straight line passing through the origin. The model just described often is called the no-intercept model.)

Question 3 (11.15 from the textbook)

1. Derive the following identity:

$$SSE = S_{yy} - \hat{\beta}_1 S_{xy}$$

   This provides an easier computational method of finding the SSE.

2. Use the computational formula for SSE derived in part (a) to prove that $SSE \leq S_{yy}$. [**Hint**: $\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$]

Question 4 (11.21 from the textbook)

Suppose that $Y_1, Y_2, ..., Y_n$ are independent normal random variables with $\mathbb{E}(Y_i) = \beta_0 + \beta_1 x_i$ and $\mathbb{V}(Y_i) = \sigma^2$, for $i = 1, 2, ..., n$. Find $Cov(\hat{\beta}_0, \hat{\beta}_1)$.
Then, prove that if $\sum_{i=1}^{n} x_i = 0$ then $\hat{\beta}_0, \hat{\beta}_1$ are independent.

Continuous Distributions

Distribution	Probability Function	Mean	Variance	Moment-Generating Function
Uniform	$f(y) = \frac{1}{\theta_2 - \theta_1};\ \theta_1 \le y \le \theta_2$	$\frac{\theta_1 + \theta_2}{2}$	$\frac{(\theta_2 - \theta_1)^2}{12}$	$\frac{e^{t\theta_2} - e^{t\theta_1}}{t(\theta_2 - \theta_1)}$
Normal	$f(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\left(\frac{1}{2\sigma^2}\right)(y-\mu)^2\right]$ $-\infty < y < +\infty$	μ	σ^2	$\exp\left(\mu t + \frac{t^2\sigma^2}{2}\right)$
Exponential	$f(y) = \frac{1}{\beta} e^{-y/\beta};\quad \beta > 0$ $0 < y < \infty$	β	β^2	$(1 - \beta t)^{-1}$
Gamma	$f(y) = \left[\frac{1}{\Gamma(\alpha)\beta^\alpha}\right] y^{\alpha-1} e^{-y/\beta};$ $0 < y < \infty$	$\alpha\beta$	$\alpha\beta^2$	$(1 - \beta t)^{-\alpha}$
Chi-square	$f(y) = \frac{(y)^{(v/2)-1} e^{-y/2}}{2^{v/2}\Gamma(v/2)};$ $y^2 > 0$	v	$2v$	$(1-2t)^{-v/2}$
Beta	$f(y) = \left[\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\right] y^{\alpha-1}(1-y)^{\beta-1};$ $0 < y < 1$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	does not exist in closed form

Discrete Distributions

Distribution	Probability Function	Mean	Variance	Moment-Generating Function
Binomial	$p(y) = \binom{n}{y} p^y (1-p)^{n-y};$ $y = 0, 1, \ldots, n$	np	$np(1-p)$	$[pe^t + (1-p)]^n$
Geometric	$p(y) = p(1-p)^{y-1};$ $y = 1, 2, \ldots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$
Hypergeometric	$p(y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}};$ $y = 0, 1, \ldots, n$ if $n \le r$, $y = 0, 1, \ldots, r$ if $n > r$	$\frac{nr}{N}$	$n\left(\frac{r}{N}\right)\left(\frac{N-r}{N}\right)\left(\frac{N-n}{N-1}\right)$	
Poisson	$p(y) = \frac{\lambda^y e^{-\lambda}}{y!};$ $y = 0, 1, 2, \ldots$	λ	λ	$\exp[\lambda(e^t - 1)]$
Negative binomial	$p(y) = \binom{y-1}{r-1} p^r (1-p)^{y-r};$ $y = r, r+1, \ldots$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$	$\left[\frac{pe^t}{1-(1-p)e^t}\right]^r$