Stat 135 Lab 13

GSI: Yutong Wang

May 1, 2020

Questions?


TOYAMA REGIONAL RAILWAY
TOYAMA REGIONAL RAILWAY
CHITETSU
CHITETSU
7022
豆瓣 @小野二

To-do Today

1. Interpretation of linear model output in R
   [See the Rmd demo]
2. SE of the regression line
3. Prediction Interval
4. Bayesian statistics
5. Practice problems

Berkeley
UNIVERSITY OF CALIFORNIA

SE of the regression line

Note: $\text{Cov}(aX + bY, cW + dV) = ac\text{Cov}(X, W) + ad\text{Cov}(X, V) + bc\text{Cov}(Y, W) + bd\text{Cov}(Y, V)$

$$\text{Lemma:} \quad \text{Cov}(\bar{y}, \hat{\beta}_1) = \text{Cov}(\hat{\beta}_0 + \hat{\beta}_1\bar{x}, \hat{\beta}_1) = \text{Cov}(\hat{\beta}_0, \hat{\beta}_1) + \bar{x}\text{Var}(\hat{\beta}_1)$$

$$= -\frac{\sigma^2\bar{x}}{n\text{Var}(x)} + \frac{\sigma^2\bar{x}}{n\text{Var}(x)} = 0$$

$$\hat{y}_i = \bar{y} + \hat{\beta}_1(x_i - \bar{x})$$

$$\Longrightarrow \text{Var}(\hat{y}_i) = \text{Var}(\bar{y} + \hat{\beta}_1(x_i - \bar{x}))$$

$$= \text{Var}(\bar{y}) + (x_i - \bar{x})^2\text{Var}(\hat{\beta}_1) + 2(x_i - \bar{x})\text{Cov}(\bar{y}, \hat{\beta}_1)$$

$$= \text{Var}(\bar{y}) + (x_i - \bar{x})^2\text{Var}(\hat{\beta}_1) = \sigma^2(\frac{1}{n} + \frac{(x_i - \bar{x})^2}{n\text{Var}(x)})$$

$$\Longrightarrow \hat{y} \sim N(\beta_0 + \beta_1 x, \sigma^2(\frac{1}{n} + \frac{(x - \bar{x})^2}{n\text{Var}(x)}))$$

$$\text{Standard error:} \quad s_{\hat{y}} = \hat{\sigma}^2(\frac{1}{n} + \frac{(x - \bar{x})^2}{n\text{Var}(x)})$$

Berkeley
UNIVERSITY OF CALIFORNIA

Prediction interval

$$\hat{y}_{n+1} = \hat{\beta}_0 + \hat{\beta}_1 x_{n+1}$$

$$\begin{aligned}
\mathrm{Var}(y_{n+1} - \hat{y}_{n+1}) &= \mathrm{Var}(y_{n+1}) + \mathrm{Var}(\hat{y}_{n+1}) \\
&= \sigma^2 + \sigma^2\left(\frac{1}{n} + \frac{(x-\bar{x})^2}{n\mathrm{Var}(x)}\right) \\
&= \sigma^2\left(1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{n\mathrm{Var}(x)}\right)
\end{aligned}$$

$$\implies y_{n+1} - \hat{y}_{n+1} \sim N\left(0, \sigma^2\left(1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{n\mathrm{Var}(x)}\right)\right)$$

$$\implies \frac{y_{n+1} - \hat{y}_{n+1} - 0}{\hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{(x_{n+1}-\bar{x})^2}{n\mathrm{Var}(x)}}} \sim t_{n-2}$$

$$\implies \mathbb{P}\left[y_{n+1} \in \hat{y}_{n+1} \pm t_{n-2}(0.025)\hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{n\mathrm{Var}(x)}}\right] = 0.95$$

Berkeley
UNIVERSITY OF CALIFORNIA

Bayesian statistics

Bayes Rule:

$$f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{f_{X|Y}(x \mid y) \cdot f_Y(y)}{f_X(x)}$$

Posterior Probability:

$$f_{\Theta|X}(\theta \mid x) = \frac{f_{X|\Theta}(x \mid \theta) f_\Theta(\theta)}{f_X(x)}$$

where $f_{\Theta|X}(\theta \mid x)$ is the posterior probability, $f_\Theta(\theta)$ is the prior probability, $f_{X|\Theta}(x \mid \theta)$ is the likelihood density, and $f_X(x)$ is a normalizing constant.

$$f_{\Theta|X}(\theta \mid x) \propto f_{X|\Theta}(x \mid \theta) f_\Theta(\theta)$$

- We treat the parameter θ as a random variable in Bayesian statistics, when the frequentists treat the parameter as some fixed constant.
- An estimate of θ is called the posteior mean $\hat{\theta} = \mathbb{E}(\Theta \mid X)$, which is a function of X.
- When the prior and posterior distributions belong to the same distribution family, we say that the prior and likelihood are conjugate.

Berkeley
UNIVERSITY OF CALIFORNIA

Bayesian statistics

Example : Beta distribution $X \sim \text{Beta}(\gamma, s)$

$$f(x) = \frac{\Gamma(\gamma+s)}{\Gamma(\gamma)\Gamma(s)} x^{\gamma-1}(1-x)^{s-1}, \mathbb{E}(x) = \frac{\gamma}{\gamma+s}$$

Denote the prior as $p \sim \text{Beta}(\gamma, s)$, and the likelihood density is $X \mid p \sim$Binomial(n, p). Then, the posterior is

$$p \mid X = x \sim \text{Beta}(x+\gamma, n-x+s)$$

and the posterior mean is

$$\hat{p} = \mathbb{E}(p \mid X = x) = \frac{x+\gamma}{n+\gamma+s} = \frac{n}{n+\gamma+s}\left(\frac{x}{n}\right) + \frac{\gamma+s}{n+\gamma+s}\left(\frac{\gamma}{\gamma+s}\right)$$

where $\frac{x}{n}$ is MLE. and $\frac{\gamma}{\gamma+s}$ is the prior mean.

Berkeley
UNIVERSITY OF CALIFORNIA

Problem 1: Bayesian Statistics

Consider a biased coin with probability of landing heads equal to θ. Also, let X be a random variable that is equal to 1 when the coin lands heads, and is otherwise equal to 0. In other words, we assume that $P(X = 1|\Theta = \theta) = \theta$, while $P(X = 0|\Theta = \theta) = 1 - \theta$. We consider the following prior for Θ :

$$f_\Theta(\theta) = \begin{cases} \theta e^\theta & \text{for } 0 \leq \theta \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the probability that a coin toss results in heads: i.e., compute $P(X = 1)$.

(b) Given that a coin toss results in heads, find the posterior density for Θ.

(c) Given that the first toss resulted in heads, find the conditional probability of heads on the second toss. (*Hint:* this is tricky. Use the posterior from part (b) as the new prior!)

Problem 2: MLE, MoM

Suppose that we have n i.i.d random variables $X_1, X_2, \ldots, X_n$, each with the following probability function (as usual, $0 \leq \theta \leq 1$ is unknown):

$$P(X=0|\theta) = \frac{2}{3}\theta, \qquad P(X=1|\theta) = \frac{1}{3}\theta, \qquad P(X=2|\theta) = \frac{2}{3}(1-\theta), \qquad P(X=3|\theta) = \frac{1}{3}(1-\theta)$$

(a) Find $\hat{\theta}_{\mathrm{ML}}$, i.e. the max likelihood estimate of θ. Also find $E(\hat{\theta}_{\mathrm{ML}})$ and $Var(\hat{\theta}_{\mathrm{ML}})$.

(b) Find $\hat{\theta}_{\mathrm{MM}}$, i.e. the method of moments estimate of θ (it will depend only on $\hat{\mu}_1 = \bar{X}$)

(c) Compute the standard error of $\hat{\theta}_{\mathrm{MM}}$. Note that it is a function of θ.

(d) How does the standard error of $\sigma_{\hat{\theta}_{\mathrm{MM}}}$ compare to the standard error of $\sigma_{\hat{\theta}_{\mathrm{ML}}}$ for different values of θ? *Hint:* It is in fact a bit simpler to compare their squares.

(e) Finally, assume that we have a sample $x = (3, 0, 2, 1, 3, 2, 1, 0, 2, 1)$, compute (i): the MLE $\hat{\theta}_{\mathrm{ML}}$ (it is a number!) and its *estimated* standard error; and (ii): the MOM $\hat{\theta}_{\mathrm{MM}}$ and its estimated standard error.

Berkeley
UNIVERSITY OF CALIFORNIA

Problem 3: Posterior Mean

Let's assume the same setup from problem 2 above but turn to the Bayeesian perspective with the flat prior $\Theta \sim Unif(0,1)$.

Find the posterior distribution of $\Theta | X_1, \ldots, X_n$.

Show that the posterior mean $E(\Theta | X_1, \ldots, X_n)$ is a weighted average of $\hat{\Theta}_{ML}$ from problem 2 and the prior mean.

Berkeley
UNIVERSITY OF CALIFORNIA