PSTAT 126 Hints for Homework I Jiaye Xu

Fall 2020

HW1 P1

1. By Def. $Var(Z) = E[(Z - E(Z))^2]$

$$= E[Z^2 - 2Z\,E(Z) + (E(Z))^2]$$

$$= E[Z^2] - [E(Z)]^2$$

$$LHS = E[(aZ + bZ')^2] - [E(aZ + b(Z'))]^2$$

[illegible]

By Def. $Cov(Z, Z') = E(ZZ') - E(Z)E(Z')$

$$= a^2\,Var(Z) + b^2\,Var(Z') + 2ab\,cov(Z, Z')$$

$$= RHS$$

2. By Def. A collection of events $A_1, \dots, A_n$ are <u>mutually independent</u> if for any subcollection $A_{i_1}, \dots, A_{i_k}$, we have

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} P(A_{i_j})$$

$A_1, \dots, A_n$ are <u>pairwise independent</u> if any two events are independent, i.e. for any $i \neq j$

$$P(A_i \cap A_j) = P(A_i)\,P(A_j)$$

N.B. mutually independence is a stronger definition

(a) pairwise independence does NOT imply mutual independence. Give one counter example.
(NOTE: Use Your own example!)

PSTAT 126
Fall 2020

Jiaye Xu

HW1

P2

Counter example:

Let sample space $S = \{ 3!$ permutation of $a, b, c \}$

i.e. $S = \begin{Bmatrix} aaa & bbb & ccc \\ abc & bca & cba \\ acb & bac & cab \end{Bmatrix}$

Define $A_i = \{ i$th place in the triple is occupied by $a \}$.

Then, $P(A_i) = \frac{1}{3}$, for $i = 1, 2, 3$

and $P(A_1 \cap A_2) = P(A_1 \cap A_3) = P(A_2 \cap A_3) = \frac{1}{9}$

Thus, A_1, A_2, A_3 are pairwise independent.

However,

$P(A_1 \cap A_2 \cap A_3) = \frac{1}{9} \neq P(A_1) P(A_2) P(A_3)$

So, A_1, A_2, A_3 are NOT mutually independent.

(b) Mutual independence implies pairwise independence.

By Def of mutual independence, for any sub collection $\{ A_{i_1}, A_{i_2} \}$, we have

$P(A_{i_1}, A_{i_2}) = P(A_{i_1}) P(A_{i_2})$

3. Yes. Because by "linear", we mean we have linearity in parameters, i.e., β_j's, not necessarily in covariates.

(NOTE: You should explain more details, and it's better to give an example.)

PSTAT 126

Fall 2020

Jiaye Xu

P3

4. Residuals $e_i = Y_i - \hat{Y}_i$, where $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ is the fitted value in a SLR

$$\sum_{i=1}^{n} e_i = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)$$

$$= \sum_{i=1}^{n} (Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)$$

[illegible]] b/c $\bar{Y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{X}$

= [illegible]

$= 0$ b/c $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$, $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

5. · Our goal of regression: $Y = f(x) + \varepsilon$.
where $f(x)$ is the deterministic component, and is an approximate functional relationship between X and Y

· Mean function $E(Y|x)$ is important in regression, because (1) we want to minimize $E[(f(x) - Y)^2] \geq 0$
and (2) we have $E[(f(x) - Y)^2]$
$$= E[(f(x) - E(Y|x))^2] + E[(Y - E(Y|x))^2]$$
By taking $f(x) = E(Y|X)$, we can minimize LHS of the equation.

This implies the mean function $E(Y|X)$ is the best representation of the functional relationship between X and Y (recall $Y \approx f(x)$)

(For details, see P6 to P7 in Lecture 2 slides)
(NOTE: Explain in your own way!)

PSTAT 126
Fall 2020

Jiaye Xu
P4

6. N.B. Use calculus, rather than derive formulae for $\hat{\beta}_0, \hat{\beta}_1$

Goal: $(\hat{\beta}_0, \hat{\beta}_1) = \arg\min_{(\beta_0, \beta_1) \in R^2} \sum_{n=1}^{N} (y_n - (\beta_0 + \beta_1 x_n))^2$

Let $f(\beta_0, \beta_1) = \sum_{n=1}^{N} (y_n - (\beta_0 + \beta_1 x_n))^2$

$$\begin{cases} \dfrac{\partial f(\beta_0, \beta_1)}{\partial \beta_0} = -2 \cdot \sum_{n=1}^{N} (y_n - (\beta_0 + \beta_1 x_n)) = 0 & (1) \\ \dfrac{\partial f(\beta_0, \beta_1)}{\partial \beta_0} = -2 \sum_{n=1}^{N} (y_n - (\beta_0 + \beta_1 x_n)) \cdot x_n = 0 & (2) \end{cases}$$

In this case $N=3$, three data points: $(0,1), (1,0), (1,1)$

$$\begin{cases} -2 \left[(1 - \beta_0) - (\beta_0 + \beta_1) + 1 - (\beta_0 + \beta_1) \right] = 0 \\ -2 \left[0 - (\beta_0 + \beta_1) + 1 - (\beta_0 + \beta_1) \right] = 0 \end{cases}$$

[redacted]

$\therefore (\hat{\beta}_0, \hat{\beta}_1) =$ [redacted]

7. you can use R example on page 12 of Lecture 2 slides.