Stats100A **Summer 2024**

Week 6: Extra Problems - Solutions

Author: Andrew Lizarraga

Solutions

6.1 Linear Regression

We are given that $X \sim N(0,1)$ and that Y is linear function of X given by $Y = \rho X + \varepsilon$, where $|\rho| < 1$ and $\varepsilon \sim N(0, 1-\rho^2)$. We are also given that $\varepsilon \perp X$.

Problem 1: Calculate $E[Y|X = x]$.

Solution: Notice that by definition of Y and conditional expectation, we have:

$$E[Y|X = x] = E[\rho X + \varepsilon|X = x] = \rho x + E[\varepsilon|X = x]$$

Now since $\varepsilon \perp X$, we have that $E[\varepsilon|X = x] = E[\varepsilon] = 0$. Therefore:

$$E[Y|X = x] = \rho x$$

Problem 2: Calculate $\text{Var}[Y|X = x]$.

Solution: Again to deduce this result, we use the fact that Y is linear in X and properties of variance and independence:

$$\text{Var}[Y|X = x] = \text{Var}[\rho X + \varepsilon|X = x]$$

Now since $X \perp \epsilon$:

$$\text{Var}[Y|X = x] = \text{Var}[\rho X|X = x] + \text{Var}[\varepsilon|X = x]$$

Note that since ρX is deterministic given X, we have: $\text{Var}[\rho X|X = x] = 0$. Additionally, since $X \perp \varepsilon$, we have $\text{Var}[\varepsilon|X = x] = \text{Var}[\varepsilon] = 1 - \rho^2$.

Problem 3: Calculate the joint density $f(x, y)$ based on the chain rule $f(x, y) = f(x)f(y|x)$.

Solution: Recall that for $X \sim N(0, 1)$ we have:

$$f(X) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$$

Also since Normal distributions have a closed form (this is from the lectures), and we computed the mean and variance of $[Y|X = x]$from problems 1 and 2, thus $[Y|X = x] \sim N(\rho x, 1 - \rho^2)$ and,

$$f(y|x) = \frac{1}{\sqrt{2\pi(1 - \rho^2)}} e^{\frac{-(y-\rho x)^2}{2(1-\rho^2)}}$$

Now by the chain rule we have:

$$f(x, y) = f(x)f(y|x) = \left(\frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} \right) \left(\frac{1}{\sqrt{2\pi(1 - \rho^2)}} e^{\frac{-(y-\rho x)^2}{2(1-\rho^2)}} \right)$$

Simplifying yields:

$$f(x, y) = \frac{1}{2\pi\sqrt{1 - \rho^2}} e^{-\frac{x^2}{2} - \frac{(y-\rho x)^2}{2(1-\rho^2)}}$$

Problem 4: Calculate $E[Y]$

Solution: The calculation is direct:

$$E[Y] = E[\rho X + \varepsilon] = \rho E[X] + E[\varepsilon]$$

Now since $E[X] = 0$ and $E[\varepsilon] = 0$, thus $E[Y] = 0$.

Problem 5: Calculate $\text{Var}[Y]$

Solution: The calculation is direct and follows from definition of Y, variance and independence of X and ε:

$$\text{Var}[Y] = \text{Var}[\rho X + \varepsilon] = \rho^2 \text{Var}[X] + \text{Var}[\varepsilon]$$

Since $\text{Var}[X] = 1$ and $\text{Var}[\epsilon] = 1 - \rho^2$, therefore:

$$\text{Var}[Y] = \rho^2 + (1 - \rho^2) = 1$$

Problem 6: Calculate $\text{Cov}(X, Y)$

Solution: Using the following formula from lecture:

$$\text{Cov}(X, Y) = E[XY] - E[X]E[Y]$$

Now we've have $E[X] = 0, E[Y] = 0$, thus

$$\text{Cov}(X, Y) = E[XY]$$

Since $Y = \rho X + \varepsilon$, thus $XY = \rho X^2 + X\varepsilon$. So we have that

$$E[XY] = E[\rho X^2] + E[X\varepsilon]$$

Since $X \perp \varepsilon$, thus $E[X\varepsilon] = E[X]E[\varepsilon] = 0$. Now we have:

$$E[XY] = \rho E[X^2]$$

Also, since $X \sim N(0, 1)$ we have $E[X^2] = \text{Var}[X] = 1$, therefore:

$$\text{Cov}(X, Y) = E[XY] = \rho$$

6.2 Transformation of Random Variables

Let $U \sim Unif[0,1]$ and let $X = -\log U$.

Problem 1: Calculate the cumulative density function $F(x) = P(X \leq x)$.

Solution: This follows directly by definition of X and monotonicity:

$$F(x) = P(X \leq x)$$

Now by definition of X:

$$= P(-\log U \leq x)$$
$$= P(\log U \geq -x)$$

Since the exponential is monotonic and convex:

$$= P(U \geq e^{-x})$$
$$= 1 - P(U \leq e^{-x})$$

Since U is a uniform random variable the probability that it's bounded between 0 and e^{-x} is just the length of this interval, namely: $P(U \leq e^{-x}) = e^{-x}$. Therefore:

$$F(x) = 1 - e^{-x}$$

Problem 2: Calculate the probability density function $f(x) = F'(x)$.

Solution:
Since $F(x) = 1 - e^{-x}$, we have that $F'(x) = e^{-x}$.