hw1

1

Prove each of the following. You may use the results from previous parts of the problem in further proofs. For each proof, I have indicated where you should start. For the purposes of these proofs here, assume that X, Y , and Z are continuous random variables. Assume that a, b and c are constants.

a

$$\begin{aligned}
E[aX + bY + c] &= \int\int (ax + by + c)p(x, y)dxdy \\
&= \int\int axp(x, y)dxdy + \int\int byp(x, y)dxdy + \int\int cp(x, y)dxdy \\
&= a\int x\Big[\int p(x, y)dy\Big]dx + b\int y\Big[\int p(x, y)dx\Big]dy + c\int\int p(x, y)dxdy \\
&= a\int xp(x)dx + b\int yp(y)dy + c \cdot 1 \\
&= aE[X] + bE[Y] + c
\end{aligned}$$

b

$$\begin{aligned}
cov(X, Y) &= E\Big[(X - E[X])(Y - E[Y])\Big] \\
&= E\Big[XY - XE[Y] - YE[X] + E[Y]E[X]\Big] \\
&= E[XY] - E[XE[Y]] - E[YE[X]] + E[E[Y]E[X]] \\
&= E[XY] - E[Y]E[X] - E[X]E[Y] + E[Y]E[X]E[1] \qquad \text{since E[X] and E[Y] are just constants} \\
&= E[XY] - E[X]E[Y]
\end{aligned}$$

c

$$\begin{aligned}
E[Y|X=x] &= \int yp(y|X=x)dy \\
&= \int y\frac{p(y,x)}{p(x)}dy \\
&= \int\int y\frac{p(y,x,z)}{p(x)}dzdy && \text{reverse marginalization} \\
&= \int\int y\frac{p(y|x,z)p(x,z)}{p(x)}dzdy && \text{bayes rule} \\
&= \int \frac{p(x,z)}{p(x)}\Big[\int yp(y|x,z)dy\Big]dz && \text{separating terms} \\
&= \int p(z|x)E[Y|X=x,Z=z]dz \\
&= E\Big[E[Y|X,Z]\;\Big|\;X=x\Big]
\end{aligned}$$

The transition between the second to last and the last step relies on the fact that since we are conditioning on $X = x$, E[Y|X,Z] is just a function of Z, meaning that the outer expectation is over $p(z|x)$.

d

$$\begin{aligned}
E[h(X,Y)] &= \int\int h(x,y)p(x,y)dxdy \\
&= \int\int h(x,y)p(y|x)p(x)dxdy && \text{bayes rule} \\
&= \int p(x)\Big[\int h(x,y)p(y|x)dy\Big]dx && \text{rearranging terms} \\
&= \int p(x)E[h(x,y)|X=x] && \text{since we condition on x, E[h(x,y)|x] is over p(y|x)} \\
&= E\Big[E[h(X,Y)|X=x]\Big] && \text{this outer expectation is just over p(x)}
\end{aligned}$$

e

$$
\begin{aligned}
Var(Y) &= E[Y^2] - E[Y]^2 \\
&= E\Big[E[Y^2|X]\Big] - E[Y]^2 \\
&= E\Big[Var(Y|X) + E[Y|X]^2\Big] - E\Big[E[Y|X]\Big]^2 \\
&= E[Var(Y|X)] + E\Big[E[Y|X]^2\Big] - E\Big[E[Y|X]\Big]^2 \\
&= E[Var(Y|X)] + Var(E[Y|X]) \qquad \text{defintion of variance}
\end{aligned}
$$

2

Prove each of the following statements. In each, let $X_{(n)}$ and $Y_{(n)}$ be sequences of random variables, and let c and d be constants.

a

If $X_{(n)} \to c$ and $Y_n \to d$ both in probability then

$$

$$

3

4

We know that $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \sigma^2)$ and $\hat{\sigma} \xrightarrow{p} \sigma$ and we can write

$$
\frac{\sqrt{n}(\hat{\theta} - \theta)}{\hat{\sigma}} = \frac{\sigma}{\hat{\sigma}} \frac{\sqrt{n}(\hat{\theta} - \theta)}{\sigma}
$$

Next we apply the continuous mapping theorem we can say that

$$
\frac{\sigma}{\hat{\sigma}} \xrightarrow{p} \frac{\sigma}{\sigma} \qquad \text{and} \qquad \frac{\sqrt{n}(\hat{\theta} - \theta)}{\sigma} \xrightarrow{d} N(0, 1)
$$

Next Slutsky's theorem tells us that

$$\frac{\sqrt{n}(\hat\theta-\theta)}{\hat\sigma} = \frac{\sigma}{\hat\sigma}\frac{\sqrt{n}(\hat\theta-\theta)}{\sigma} \xrightarrow{d} 1\cdot N(0,1)$$

And by definition of convergence in distribution we get that

$$\lim_{n\to\infty} P\bigg(- z_{1-\alpha/2} \le \frac{\sqrt{n}(\hat\theta-\theta)}{\hat\sigma} \le z_{1-\alpha/2}\bigg) = 1-\alpha$$

We can then do some rearranging within the parentheses

$$\begin{aligned}
-z_{1-\alpha/2} \le \frac{\sqrt{n}(\hat\theta-\theta)}{\hat\sigma} \le z_{1-\alpha/2} &= -z_{1-\alpha/2}\frac{\hat\sigma}{\sqrt{n}} \le \hat\theta-\theta \le z_{1-\alpha/2}\frac{\hat\sigma}{\sqrt{n}} \\
&= \hat\theta - z_{1-\alpha/2}\frac{\hat\sigma}{\sqrt{n}} \le -\theta \le \hat\theta + z_{1-\alpha/2}\frac{\hat\sigma}{\sqrt{n}} \\
&= \hat\theta - z_{1-\alpha/2}\frac{\hat\sigma}{\sqrt{n}} \le \theta \le \hat\theta + z_{1-\alpha/2}\frac{\hat\sigma}{\sqrt{n}} \qquad \text{symmetry of normal dist}
\end{aligned}$$

so we have that

$$\lim_{n\to\infty} P\bigg(\hat\theta - z_{1-\alpha/2}\frac{\hat\sigma}{\sqrt{n}} \le \theta \le \hat\theta + z_{1-\alpha/2}\frac{\hat\sigma}{\sqrt{n}}\bigg) = 1-\alpha$$

as desired.

5

a

Well in OLS we choose β such that $\mathbf{X}^T(\mathbf{Y}-\mathbf{X}\beta)=0$ ($\hat\beta_{OLS}$ solves this equation).

We can write this out as follows

$$\begin{pmatrix} 1 & \dots & 1 \\ X_{1,1} & \dots & X_{n,1} \\ \vdots & \ddots & \vdots \\ X_{1,k-1} & \dots & X_{n,k-1} \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

but immediately we can see that an immediate consequence of this is that

$$1 \cdot e_1 + 1 \cdot e_2 + ... + 1 \cdot e_n = \sum_{i=1}^{n} e_i = 0$$

so of course $\frac{1}{n} \sum e_i = 0$ as well.

b