Problem 1

Let (Z, W) be Bivariate Normal, defined as

$$\begin{aligned} Z &= X, \\ W &= \rho X + \sqrt{1-\rho^2}Y, \end{aligned}$$

with X, Y i.i.d. $\mathsf{N}(0, 1)$, and $-1 < \rho < 1$. Find $\mathsf{E}(W \mid Z)$ and $\mathrm{Var}(W \mid Z)$.

$$E(W|Z) = E(\rho X + \sqrt{1-\rho^2}\,Y \mid Z = z)$$

Condition on what we want to know, which is the event that $Z = z$

$$= E(\rho X \mid Z = z) + E(\sqrt{1-\rho^2}\,Y \mid Z = z)$$

By linearity

$$= E(\rho X \mid Z = z) + \sqrt{1-\rho^2}\cdot \underbrace{E(Y \mid Z = z)}$$

Pull constant out

Because $Z = X$, this term is equivalent to $E(Y|X)$. Given that $X \perp\!\!\!\perp Y$, and independent r.v.s convey no information about the other, we know that $E(Y|X) = E(Y)$. Given $Y \sim N(0,1)$, $E(Y) = 0$ and we can drop the term:

$$= E(\rho X \mid Z = z)$$

$$= \rho E(X \mid Z = z)$$

$$= \rho E(X \mid X = x) = \rho X = \rho Z$$

Substitution of Z for X and pulling out ρ allows this simplification.

$$\mathrm{Var}(W|Z) = E(W^2|Z) - (E(W|Z))^2$$

$$= E(W^2|Z) - (\rho Z)^2$$

$$= E((\rho X + \sqrt{1-\rho^2}\,Y)^2 \mid Z = z) - \rho^2 Z^2$$

$$= E(\rho^2 X^2 + 2\rho X\sqrt{1-\rho^2}\,Y + (1-\rho^2)Y^2 \mid Z = z) - \rho^2 Z^2$$

$$= \rho^2 \underbrace{E(X^2|Z = z)} + 2\rho\sqrt{1-\rho^2}\,\underbrace{E(XY|Z = z)} + (1-\rho^2)E(Y^2|Z = z) - \rho^2 Z^2$$

Equivalent to $E(X|Z=z)\cdot E(Y|Z=z)$ given independence and since we already know $E(X|Z=z) = 0$, this term can be dropped.

Equivalent to $\rho^2 E(X^2|X = x) = \rho^2 X^2$, which cancels with the last term

$$= (1-\rho^2)E(Y^2|Z = z)$$

Since $Z = X$ and $X \perp\!\!\!\perp Y$, $E(Y^2|X = x) = E(Y^2)$. To calculate, use the variance formula and $Y \sim N(0,1)$ to get $\mathrm{Var}(Y) = 1$.

$1 = \text{Var}(Y) = E(Y^2) - \underbrace{E(Y)^2}_{=0}$

$1 = E(Y^2)$

$$\text{Var}(W|Z) = (1-\rho^2)\cdot 1 = 1-\rho^2$$

Problem 2

Let $\mathbf{X} = (X_1, X_2, X_3, X_4, X_5) \sim \text{Mult}_5(n, \mathbf{p})$ with $\mathbf{p} = (p_1, p_2, p_3, p_4, p_5)$.

(a) Find $\text{E}(X_1 \mid X_2)$ and $\text{Var}(X_1 \mid X_2)$.

(b) Find $\text{E}(X_1 \mid X_2 + X_3)$.

(a) START BY CONDITIONING THE MULTINOMIAL ON X_2:

$$(X_1, X_3, X_4, X_5) \mid X_2 \sim \text{Mult}_5(n-n_2, p_1', p_3', p_4', p_5') \quad \text{WHERE } p_j' = \frac{p_j}{(1-p_2)}$$

THE MARGIN $X_1 | X_2 \sim \text{Bin}(n-n_2, p_1')$ AND THE EXPECTATION OF A BINOMIAL IS SIMPLY $n \cdot p$ SO:

$$E(X_1 | X_2) = (n-n_2)p_1' = \frac{(n-n_2)p_1}{(1-p_2)}$$

THE VARIANCE OF A BINOMIAL IS $np(1-p)$ AND WE GET:

$$\text{Var}(X_1 | X_2) = (n-n_2)p_1'(1-p_1') = \frac{p_1 - p_2p_1 - p_1^2}{(1-p_2)^2}(n-n_2)$$

(b) THE MULTINOMIAL LUMPING THEORY GIVES $X_2 + X_3 \sim \text{Bin}(n, p_2 + p_3)$.

AS BEFORE, CONDITION ON $X_2 + X_3$:

$$(X_1, X_4, X_5) \mid X_2 + X_3 \sim \text{Mult}_5(n-(n_2+n_3), p_1', p_4', p_5') \quad \text{WHERE } p_j' = \frac{p_j}{(1-(p_2+p_3))}$$

$X_1 | X_2 + X_3 \sim \text{Bin}(n-(n_2+n_3), p_1')$:

$$E(X_1 | X_2 + X_3) = \frac{(n-(n_2+n_3))p_1}{1-(p_2+p_3)}$$

Problem 3

Show that the following version of LOTP follows from Adam's law: for any event A and continuous random variable X with PDF f_X:

$$P(A) = \int_{-\infty}^{\infty} P(A \mid X = x) f_X(x)\, dx.$$

ADAM'S LAW PROVIDES THAT $E(E(A|X)) = E(A)$

THE FUNDAMENTAL BRIDGE PROVIDES THAT $E(I_A) = P(A)$ AND WITH CONDITIONING, $E(I_A|X) = P(A|X)$

COMBINED, WE HAVE $E(P(A|X)) = P(A)$

EXPECTATION OF CONTINUOUS R.V.S IS DEFINED $E(X) = \int_{-\infty}^{\infty} x f_X(x)\,dx$

BY SUBSTITUTING $P(A|X)$, WE GET THE STATED VERSION OF LOTP:

$$P(A) = \int_{-\infty}^{\infty} P(A|X=x) f_X(x)\,dx$$

Problem 4

Let $N \sim \text{Pois}(\lambda_1)$ be the number of movies that will be released next year. Suppose that for each movie the number of tickets sold is $\text{Pois}(\lambda_2)$, independently.

(a) Find the mean and the variance of the number of movie tickets that will be sold next year.

(b) Use simulations in R (the statistical programming language) to numerically estimate mean and the variance of the number of movie tickets that will be sold next year assuming that the mean number of movies released each year in the US is 700, and that, on average, 800000 tickets were sold for each movie.

(a) THE NUMBER OF MOVIE TICKETS SOLD NEXT YEAR IS THE NUMBER OF MOVIES RELEASED MULTIPLIED BY THE AVERAGE NUMBER OF TICKETS SOLD. IF THE NUMBER OF TICKETS SOLD FOR EACH MOVIE IS DEFINED BY $M \sim \text{Pois}(\lambda_2)$, THEN THE EXPECTED NUMBER OF MOVIE TICKETS SOLD IS:

$$E(T) = E(NM) = E(N)E(M)$$

WHERE T IS THE NUMBER OF TICKETS SOLD AND $N \perp\!\!\!\perp M$

$$= \lambda_1 \cdot \lambda_2$$

THE VARIANCE OF MOVIE TICKETS SOLD IS

$$\text{Var}(T) = E(T^2) - (ET)^2 = E((NM)^2) - E(NM)^2$$

$$= E(N^2M^2) - (\lambda_1\lambda_2)^2$$

$$= E(N^2)E(M^2) - \lambda_1^2\lambda_2^2$$

$E(N^2)$ CAN BE FOUND USING THE VARIANCE FORMULA AND THE FACT THAT IF $N \sim \text{Pois}(\lambda_1)$ THEN $\text{Var}(N) = \lambda_1$.

$$\text{Var}(N) = E(N^2) - (EN)^2$$

$$\Rightarrow \lambda_1 = E(N^2) - \lambda_1^2$$

$$\Rightarrow E(N^2) = \lambda_1 + \lambda_1^2$$

THE SAME IS TRUE FOR M SUCH THAT $E(M^2) = \lambda_2 + \lambda_2^2$

$$= (\lambda_1 + \lambda_1^2)(\lambda_2 + \lambda_2^2) - \lambda_1^2 \lambda_2^2$$

$$= \lambda_1 \lambda_2 + \lambda_1 \lambda_2^2 + \lambda_1^2 \lambda_2 + \lambda_1^2 \lambda_2^2 - \lambda_1^2 \lambda_2^2$$

$$\text{Var}(T) = \lambda_1 \lambda_2 + \lambda_1 \lambda_2^2 + \lambda_1^2 \lambda_2$$

(b) SEE RMARKDOWN

Problem 5

Show that if $\mathsf{E}(Y \mid X) = c$ is a constant, then X and Y are uncorrelated. Hint: Use Adam's law to find $\mathsf{E}(Y)$ and $\mathsf{E}(XY)$.

GRANTED

ADAM'S LAW GRANTS $E(Y) = E(E(Y|X))$

THE FORMULA FOR CORRELATION IS $\text{Corr}(X,Y) = \dfrac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$

THE FORMULA FOR COVARIANCE IS $\text{Cov}(X,Y) = E(XY) - E(X)E(Y)$

IF X AND Y ARE INDEPENDENT, THEN THEIR COVARIANCE IS 0 AND THEY ARE UNCORRELATED.

USING THESE FOUNDATIONS:

$$E(XY) = E(E(XY|X))$$

$$= E(X \cdot E(Y|X))$$ FACTOR OUT X SINCE WE'RE CONDITIONING ON X AS KNOWN

$$= E(X)E(E(Y|X))$$

$$= E(X)E(Y)$$ VIA ADAMS LAW

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

$$= E(X)E(Y) - E(X)E(Y) = 0$$ SUBSTITUTE $E(X)E(Y)$ FOR $E(XY)$

Problem 6

Show that for any random variables X and Y,

$$\mathrm{E}(Y \mid \mathrm{E}(Y \mid X)) = \mathrm{E}(Y \mid X).$$

Hint: use Adam's law with extra conditioning.

GRANTED: ADAM'S LAW WITH EXTRA CONDITIONING: $E(E(Y|X,Z)|Z) = E(Y|Z)$

LET $E(Y|X) = g(x)$

RHS $E(Y|X) = E(E(Y|X,g(x))|g(x))$ VIA AND SUBSTITUTION

$= E(E(Y|x)|g(x))$ SINCE $E(Y|X,g(x)) = E(Y|X)$

$= E(g(x)|g(x))$ SUBSTITUTING $E(Y|X) = g(x)$

$= E(Y|X)$

Problem 6

Let Y denote the number of heads in n flips of a coin, whose probability of heads is θ.

(a) Suppose θ follows a distribution $\mathsf{P}(\theta) = \mathsf{Beta}(a, b)$, and then you observe y heads out of n flips. Show algebraically that the mean $\mathsf{E}(\theta \mid Y = y)$ always lies between the mean $\mathsf{E}(\theta)$ and the observed relative frequency of heads:

$$\min\left\{\mathsf{E}(\theta), \frac{y}{n}\right\} \leq \mathsf{E}(\theta \mid Y = y) \leq \max\left\{\mathsf{E}(\theta), \frac{y}{n}\right\}.$$

Here $\mathsf{E}(\theta \mid Y = y)$ is the mean of the distribution $\mathsf{P}(\theta \mid Y = y)$, and $\mathsf{E}(\theta)$ is the mean of the distribution $\mathsf{P}(\theta) = \mathsf{Beta}(a, b)$.

(b) Show that, if θ follows a uniform distribution,

$$\mathsf{P}(\theta) = \mathsf{Unif}(0, 1),$$

we have

$$\mathsf{Var}(\theta \mid Y = y) \leq \mathsf{Var}(\theta).$$

Here $\mathsf{Var}(\theta \mid Y = y)$ is the variance of the distribution $\mathsf{P}(\theta \mid Y = y)$, and $\mathsf{Var}(\theta)$ is the variance of the distribution $\mathsf{P}(\theta) = \mathsf{Unif}(0, 1)$.

(a) $Y \mid \theta \sim \text{BIN}(n, p)$ AND WE WANT TO KNOW $P(\theta \mid Y = y)$.

$$P(\theta \mid Y = y) = \frac{P(Y \mid \theta = \theta) P(\theta)}{P(Y)} \quad \text{VIA BAYES RULE}$$

$$\binom{n}{y} \theta^{y+a-1} (1-\theta)^{n-y+b-1} \cdot \frac{1}{P(Y)} \cdot \frac{1}{\beta(a,b)} \quad \text{VIA PLUGGING THE BINOMIAL AND BETA PDFs}$$

REMOVE THE NORMALIZING CONSTANTS SINCE $P(\theta \mid Y = y) \propto P(Y \mid \theta = \theta) P(\theta)$:

$$\theta^{y+a-1} (1-\theta)^{n-y+b-1} = \beta(y+a, n-y+b)$$

THE EXPECTED VALUE OF A BETA DISTRIBUTION IS $\frac{a}{a+b}$, WHICH GIVES $\frac{y+a}{a+b-y}$

BECAUSE $E(\theta \mid Y = y) = c_1 E(\theta) + c_2 \frac{y}{n}$ WHERE $c_1, c_2 > 0$ AND $c_1 + c_2 = 1$, THE POSTERIOR $P(\theta \mid Y = y)$ IS CALCULATED BASED ON THE PRIOR AND THE NEW OBSERVATION $\frac{y}{n}$. GIVEN THIS, WE KNOW THAT $E(\theta \mid Y = y)$ MUST BE BETWEEN $\frac{y}{n}$ AND THE PRIOR $E(\theta)$. INTUITIVELY, THIS MAKES SENSE SINCE AFTER YOU SEE THE DATA, THE RESULT WILL BE BETWEEN PRIOR AND THE OBSERVATION.

(b) VARIANCE OF A $\text{UNIF}(0,1)$ IS $\frac{(b-a)^2}{12}$ SO $\text{VAR}(\theta) = \frac{1}{12}$

DUE TO THE RELATION BETWEEN THE BETA AND UNIFORM DISTRIBUTIONS, WE KNOW THAT $\text{BETA}(1,1) = \text{UNIF}(0,1)$. FROM (a), WE ALSO KNOW $P(\theta \mid Y = y) \propto P(Y \mid \theta = \theta) P(\theta)$, SO:

$$P(Y \mid \theta) P(\theta) \propto \theta^{y+1-1} (1-\theta)^{n-y+1-1} \Rightarrow \text{BETA}(y+1, n-y+1)$$

THE VARIANCE OF A BETA IS $\frac{ab}{(a+b)^2(a+b+1)}$ SO:

$$VAR(\theta|Y=y) = \frac{-y^2+ny+n+1}{(n+2)^2(n+3)}$$

TO SHOW THAT $VAR(\theta|Y=y) \leq \frac{1}{12}$, WE FIND THE CRITICAL VALUES OF THE VARIANCE:

$$VAR'(\theta|Y=y) = \frac{-2y+n}{(n+2)^2(n+3)} = 0$$, SOLVING FOR y GIVES $\frac{n}{2}$

$$VAR''(\theta|Y=y) = \frac{-2}{(n+2)^2(n+3)} \Rightarrow$$ THIS IS A NEGATIVE CONSTANT FOR ALL y.

GIVEN THIS, WE KNOW THAT $y = \frac{n}{2}$ IS THE MAXIMUM VARIANCE AT $\frac{1}{12}$

AND $VAR(\theta|Y=y) \leq \frac{1}{12}$

Problem 7

Let A, B and C be independent random variables with the following distributions:

$$P(A = 1) = 0.4,\ P(A = 2) = 0.6$$

$$P(B = -3) = 0.25,\ P(B = -2) = 0.25,\ P(B = -1) = 0.25,\ P(B = 1) = 0.25$$

$$P(C = 1) = 0.5,\ P(C = 2) = 0.4,\ P(C = 3) = 0.1$$

(a) What is the probability that the quadratic equation

$$Ax^2 + Bx + C = 0$$

has two real roots that are different?

(b) What is the probability that the quadratic equation

$$Ax^2 + Bx + C = 0$$

has two real roots that are both strictly positive?

SEE R MARKDOWN