Assignment 2, Mathematical Statistics 3

Andrew Martin

April 17, 2018

1. Suppose X_1 and X_2 are discrete random variables with $X_1 \sim B(n, \pi)$ and $X_2|X_1 = x_1 \sim B(x_1, \rho)$

   (a) Write down the joint PDF of (X_1, X_2)

   Solution Using the conditional probability function statement:

$$p_{X_2|X_1}(x_2|x_1) = \frac{p(x_1, x_2)}{p_{X_1}(x_1)}$$

   Rearranged:

$$p(x_1, x_2) = p_{X_2|X_1}(x_2|x_1)p_{X_1}(x_1) = B(n, \pi)B(x_1, \rho)$$

$$p(x_1, x_2) = \binom{n}{x_1}\pi^{x_1}(1-\pi)^{n-x_1}\binom{x_1}{x_2|x_1}\rho^{x_2|x_1}(1-\rho)^{x_1-x_2|x_1}$$

   As Required

   (b) Derive the Marginal distribution of X_2

   Solution Marginal of X_2 is sum over X_1.

$$\begin{aligned}
p_{X_2}(x_2) &= \sum_{x_1}\binom{n}{x_1}\pi^{x_1}(1-\pi)^{n-x_1}\binom{x_1}{x_2}\rho^{x_2}(1-\rho)^{x_1-x_2} \\
&= \sum_{x_1=0}^{n}\binom{n}{x_1}\pi^{x_1}(1-\pi)^{n-x_1}\binom{x_1}{x_2}\rho^{x_2}(1-\rho)^{x_1-x_2} \\
&= \rho^{x_2}\sum_{x_1=0}^{n}\frac{n!}{x_1!(n-x_1)!}\pi^{x_1}(1-\pi)^{n-x_1}\frac{x_1!}{x_2!(x_1-x_2!)}(1-\rho)^{x_1-x_2} \\
&= \rho^{x_2}\sum_{x_1=0}^{n}\pi^{x_1}(1-\pi)^{n-x_1}\frac{n!}{x_2!(n-x_1)!(x_1-x_2!)}(1-\rho)^{x_1-x_2}
\end{aligned}$$

   As Required

2. (a) Consider pairs of RVs

$$(Y_{11}, Y_{12}), (Y_{21}, Y_{22}), \ldots, (Y_{n1}, Y_{n2})$$

   such that $E(Y_{i1}) = \mu_1$ and $E(Y_{i2}) = \mu_2$, $cov(Y_{i1}, Y_{i2}) = \sigma_{12}$ and Y_{ij}, Y_{kl} are independent for $i \neq k$. If $X_1 = \sum_{i=1}^{n} Y_{i1}$ and $X_2 = \sum_{j=1}^{n} Y_{j2}$ show that $cov(X_1, X_2) = n\sigma_{12}$

Solution Note $cov(Y_{i1}, Y_{i2}) = E((Y_{i1} - E(Y_{i1}))(Y_{i2} - E(Y_{i2}))) = \sigma_{12}$

$$\begin{aligned}
cov(X_1, X_2) &= E\left((X_1 - E(X_1))(X_2 - E(X_2))\right) \\
&= E\left((\sum_{i=1}^{n} Y_{i1} - E(\sum_{i=1}^{n} Y_{i1}))(\sum_{j=1}^{n} Y_{j2} - E(\sum_{j=1}^{n} Y_{j2})) \right) \\
&= E\left((\sum_{i=1}^{n} Y_{i1} - \sum_{i=1}^{n} E(Y_{i1}))(\sum_{j=1}^{n} Y_{j2} - \sum_{j=1}^{n} E(Y_{j2})) \right) \\
&= E\left((\sum_{i=1}^{n} (Y_{i1} - E(Y_{i1})))(\sum_{j=1}^{n} (Y_{j2} - E(Y_{j2}))) \right) \\
&= E\left(\sum_{\substack{i=1 \\ j=1}}^{n} (Y_{i1} - E(Y_{i1}))(Y_{j2} - E(Y_{j2})) \right) \\
&= \sum_{\substack{i=1 \\ j=1}}^{n} (cov(Y_{i1}, Y_{j2})) \\
&\text{separate cases } i = j \text{ and } i \neq j \\
&= \sum_{i=1}^{n} cov(Y_{i1}, Y_{i2}) + \sum_{\substack{i=1 \\ j=1 \\ i \neq j}}^{n} cov(Y_{i1}, Y_{j2}) \\
&= \sum_{i=1}^{n} \sigma_{12} + 0 \text{ due to independence.} \\
&= n\sigma_{12}
\end{aligned}$$

As Required

(b) Consider an experiment which results in exactly one of:

- Outcome 1, with probability π_1,
- Outcome 2, with probability π_2,
- Outcome 3, with probability $1 - \pi_1 - \pi_2$

Let Y_1 and Y_2 be indicator variables defined by

$$Y_1 = \begin{cases} 1 & \text{for outcome 1} \\ 0 & \text{otherwise} \end{cases} \qquad and \quad Y_2 = \begin{cases} 1 & \text{for outcome 2} \\ 0 & \text{otherwise} \end{cases}$$

Show that $cov(Y_1, Y_2) = -\pi_1\pi_2$

Solution

$$\begin{aligned}
cov(Y_1, Y_2) &= E\left((Y_1 - E(Y_1))(Y_2 - E(Y_2))\right) \\
&= E(Y_1Y_2) - E(Y_1)E(Y_2) \\
&= E(Y_1Y_2) - \pi_1\pi_2 \\
&= 0 - \pi_1\pi_2 \text{ as } Y_1 = 1 \implies Y_2 = 0 \text{ and } Y_2 = 1 \implies Y_1 = 0 \\
&= -\pi_1\pi_2
\end{aligned}$$

As Required

(c) Hence show that $cov(X_1, X_2) = -n\pi_1\pi_2$ if (X_1, X_2) have the trinomial distribution with parameters n and π_1, π_2

Solution Using a, $cov(X_1, X_2) = n\sigma_{12}$, where $\sigma_{12} = cov(Y_1, Y_2)$. So in this case, $\sigma_{12} = -\pi_1\pi_2$. Which results in:

$$cov(X_1, X_2) = -n\pi_1\pi_2$$

As Required

3. Suppose X_1, X_2 have joint PDF

$$f(x_1, x_2) = k(x_1 + x_2^2), \text{ for } 0 \leq x_1, x_2 \leq 1$$

(a) Find the value of k for which $f(x_1, x_2)$ is a valid PDF

Solution f is a valid PDF if $f(x_1,x_2) \geq 0, \forall x_1, x_2$ and $\int_0^1 \int_0^1 f(x_1,x_2)dx_1dx_2 = 1$

$$\begin{aligned}
\int_0^1 \int_0^1 f(x_1,x_2)dx_1dx_2 &= \int_0^1 \int_0^1 k(x_1 + x_2^2)dx_1dx_2 \\
&= k\int_0^1 (\frac{x_1^2}{2} + x_1x_2^2)|_{x_1=0}^{x_1=1}dx_2 \\
&= k\int_0^1 (\frac{1}{2} + x_2^2)dx_2 \\
&= k\left(\frac{x_2}{2} + \frac{x_2^3}{3}\right)|_{x_2=0}^{x_2=1} \\
&= k\left(\frac{1}{2} + \frac{1}{3}\right) \\
&= \frac{5}{6}k = 1 \\
&\implies k = \frac{6}{5}
\end{aligned}$$

As Required

(b) Find $P(X_1 > X_2)$

Solution

$$\begin{aligned}
\int_0^1 \int_0^{x_1} f(x_1,x_2)dx_2dx_1 &= \int_0^1 \int_0^{x_1} \frac{6}{5}(x_1 + x_2^2)dx_2dx_1 \\
&= \frac{6}{5}\int_0^1 (x_1x_2 + \frac{x_2^3}{3})|_{x_2=0}^{x_2=x_1}dx_1 \\
&= \frac{6}{5}\int_0^1 (x_1^2 + \frac{x_1^3}{3})dx_1 \\
&= \frac{6}{5}(\frac{x_1^3}{3} + \frac{x_1^4}{12})|_0^1 \\
&= \frac{6}{5}(\frac{1}{3} + \frac{1}{12}) = \frac{1}{2}
\end{aligned}$$

As Required

(c) Find $P(X_1 + X_2 \leq \frac{1}{2})$

Solution

$$\begin{aligned}
\int_0^{\frac{1}{2}} \int_0^{\frac{1}{2}-x_1} f(x_1,x_2)dx_2dx_1 &= \int_0^{\frac{1}{2}} \int_0^{\frac{1}{2}-x_1} \frac{6}{5}(x_1 + x_2^2)dx_2dx_1 \\
&= \frac{6}{5}\int_0^{\frac{1}{2}} (x_1x_2 + \frac{x_2^3}{3})|_{x_2=0}^{x_2=\frac{1}{2}-x_1}dx_1 \\
&= \frac{6}{5}\int_0^{\frac{1}{2}} (x_1(\frac{1}{2} - x_1) + \frac{(\frac{1}{2}-x_1)^3}{3})dx_1 \\
&= \frac{6}{5}\int_0^{\frac{1}{2}} (\frac{x_1}{2} - x_1^2 + \frac{(\frac{1}{2}-x_1)^3}{3})dx_1 \\
&= \frac{6}{5}(\frac{x_1^2}{4} - \frac{x_1^3}{3} + \frac{(\frac{1}{2}-x_1)^4}{-12})|_0^{\frac{1}{2}} \\
&= \frac{6}{5}(\frac{1}{16} - \frac{1}{24} - \frac{\frac{1}{2^4}}{-12}) = \frac{1}{32}
\end{aligned}$$

As Required

(d) Find $P(X_1 \leq \frac{1}{4})$

Solution

$$\int_0^{\frac{1}{4}}\int_0^1 f(x_1,x_2)dx_2dx_1 = \int_0^{\frac{1}{4}}\int_0^1 \frac{6}{5}(x_1+x_2^2)dx_2dx_1$$
$$= \frac{6}{5}\int_0^{\frac{1}{4}}(x_1x_2+\frac{x_2^3}{3})|_{x_2=0}^{x_2=1}dx_1$$
$$= \frac{6}{5}\int_0^{\frac{1}{4}}(x_1+\frac{1}{3})dx_1$$
$$= \frac{6}{5}(\frac{x_1^2}{2}+\frac{1}{3}x_1)|_0^{\frac{1}{4}}$$
$$= \frac{6}{5}(\frac{1}{4^2*2}+\frac{1}{3}\frac{1}{4})$$
$$= \frac{6}{5}\frac{1}{32}+\frac{6}{5}\frac{1}{12} = \frac{11}{80}$$

As Required

4. Suppose (X_1, X_2) have the Dirichlet distribution:

$$f(x_1,x_2) = \frac{\Gamma(\alpha_1+\alpha_2+\alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)}x_1^{\alpha_1-1}x_2^{\alpha_2-1}(1-x_1-x_2)^{\alpha_3-1}$$

(a) Prove that the marginal distribution of X_1 is a Beta distribution

Solution Recall Beta(α,β) distribution has form:

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$$

Marginal of X_1:

$$f_{X_1}(x_1) = \int_0^{1-x_1}\frac{\Gamma(\alpha_1+\alpha_2+\alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)}x_1^{\alpha_1-1}x_2^{\alpha_2-1}(1-x_1-x_2)^{\alpha_3-1}dx_2$$
$$= \frac{\Gamma(\alpha_1+\alpha_2+\alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)}x_1^{\alpha_1-1}\int_0^{1-x_1}x_2^{\alpha_2-1}(1-x_1-x_2)^{\alpha_3-1}dx_2$$

let $x_2 = (1-x_1)t \implies dx_2 = dt$

$$= \frac{\Gamma(\alpha_1+\alpha_2+\alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)}x_1^{\alpha_1-1}\int_0^1((1-x_1)t)^{\alpha_2-1}(1-t)^{\alpha_3-1}dt$$
$$= \frac{\Gamma(\alpha_1+\alpha_2+\alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)}x_1^{\alpha_1-1}(1-x_1)^{\alpha_2+\alpha_3-1}\int_0^1 t^{\alpha_2-1}(1-t)^{\alpha_3-1}dt$$

this integral is the beta function $Beta(\alpha_2,\alpha_3) = \frac{\Gamma(\alpha_2)\Gamma(\alpha_3)}{\Gamma(\alpha_2+\alpha_3)}$

$$= \frac{\Gamma(\alpha_1+\alpha_2+\alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)}\frac{\Gamma(\alpha_2)\Gamma(\alpha_3)}{\Gamma(\alpha_2+\alpha_3)}x_1^{\alpha_1-1}(1-x_1)^{\alpha_2+\alpha_3-1}$$
$$= \frac{\Gamma(\alpha_1+\alpha_2+\alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2+\alpha_3)}x_1^{\alpha_1-1}(1-x_1)^{\alpha_2+\alpha_3-1}$$

which is the beta distribution: $Beta(\alpha_1,\alpha_2+\alpha_3)$

As Required

(b) Find the conditional density function $f_{X_2|X_1}(x_2|x_1)$

Solution Using the conditional density statement:

$$f_{X_2|X_1}(x_2|x_1) = \frac{f(x_1,x_2)}{f_{X_1}(x_1)}$$
$$= \left(\frac{\Gamma(\alpha_1+\alpha_2+\alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)}x_1^{\alpha_1-1}x_2^{\alpha_2-1}(1-x_1-x_2)^{\alpha_3-1}\right)/\left(\frac{\Gamma(\alpha_1+\alpha_2+\alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2+\alpha_3)}x_1^{\alpha_1-1}(1-x_1)^{\alpha_2+\alpha_3-1}\right)$$
$$= \left(\frac{x_2^{\alpha_2-1}(1-x_1-x_2)^{\alpha_3-1}}{\Gamma(\alpha_2)\Gamma(\alpha_3)}\right)/\left(\frac{(1-x_1)^{\alpha_2+\alpha_3-1}}{\Gamma(\alpha_2+\alpha_3)}\right)$$
$$= \frac{\Gamma(\alpha_2+\alpha_3)}{\Gamma(\alpha_2)\Gamma(\alpha_3)}\frac{x_2^{\alpha_2-1}(1-x_1-x_2)^{\alpha_3-1}}{(1-x_1)^{\alpha_2+\alpha_3-1}}$$
$$= \frac{\Gamma(\alpha_2+\alpha_3)}{\Gamma(\alpha_2)\Gamma(\alpha_3)}x_2^{\alpha_2-1}(1-x_1-x_2)^{\alpha_3-1}(1-x_1)^{1-\alpha_2-\alpha_3}$$

As Required

5. Suppose X_1, X_2 have the uniform distribution on the region $|x_1| + |x_2| \leq 1$

(a) Given an expression for the joint PDF (Sketch the region $|x_1| + |x_2| \leq 1$)
Solution This region is a diamond centred at the origin. This is effectively a rotated square with side length $\sqrt{2}$. The area of this region will be $\sqrt{2}^2 = 2$. Since it is uniform, the valid PDF would be:

$$f(x_1, x_2) = \begin{cases} \frac{1}{2}, & \text{for } |x_1| + |x_2| \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

So the regions for x_1 and x_2 which this is valid are

$$-1 + |x_2| < x_1 < 1 - |x_2| \quad and \quad -1 + |x_1| < x_2 < 1 - |x_1|$$

As Required

(b) Find $E(X_1)$ and $E(X_2)$
Solution

$$\begin{aligned} E(X_1) &= \int_{X_2} \int_{X_1} x_1 \frac{1}{2} dx_1 dx_2 \\ &= \int_{X_2} \frac{x_1^2}{4} \Big|_{-1+|x_2|}^{1-|x_2|} dx_2 \\ &= \int_{X_2} \frac{(1 - |x_2|)^2 - (-1 + |x_2|)^2}{4} dx_2 \\ &= \int_{-1}^{1} 0 dx_2 \\ &= 0 \end{aligned}$$

Similarly $E(X_2) = 0$
As Required

(c) Find $cov(X_1, X_2)$
Solution

$$\begin{aligned} Cov(X_1, X_2) &= E\left((X_1 - E(X_1))(X_2 - E(X_2))\right) \\ &= E(X_1 X_2) \\ &= \int_{-1}^{1} \int_{-1+|x_2|}^{1-|x_2|} x_1 x_2 \frac{1}{2} dx_1 dx_2 \\ &= \int_{-1}^{1} \frac{x_1^2 x_2}{4} \Big|_{-1+|x_2|}^{1-|x_2|} dx_2 \\ &= \int_{-1}^{1} \frac{x_2}{4} \left((1 - |x_2|)^2 - (-1 + |x_2|)^2\right) dx_2 \\ &= \int_{-1}^{1} 0 dx_2 \\ &= 0 \end{aligned}$$

As Required

(d) Find the marginal distribution of X_1 and also of X_2
Solution

$$\begin{aligned} f_{X_1}(x_1) &= \int_{X_2} \frac{1}{2} dx_2 \\ &= \frac{x_2}{2} \Big|_{-1+|x_1|}^{1-|x_1|} \\ &= 1 - |x_1| \end{aligned}$$

Likewise, $f_{X_2}(x_2) = 1 - |x_2|$.
As Required

(e) Are X_1 and X_2 independent? Comment on this example
Solution They are independent if the joint pdf $f(x_1, x_2) = f_{X_1}(x_1) \times f_{X_2}(x_1)$
They are not independent. As this does not hold in this case. **As Required**

6. Suppose $U \sim U(0, 1)$ and $V|u \sim U(0, u)$

(a) Write down the joint probability density function of (U, V) including its domain

Solution

$$f_U(u) = \begin{cases} 1 & 0 < u < 1 \\ 0 & \text{otherwise} \end{cases} \qquad f_{V|u}(v|u) = \begin{cases} \frac{1}{u} & 0 < v < u \\ 0 & \text{otherwise} \end{cases}$$

Again using the conditional formula

$$f(x_1, x_2) = f_{X_2|X_1}(x_2|x_1) f_{X_1}(x_1)$$

$$\begin{aligned} f(u, v) &= f_{v|u}(v|u) f_u(u) \\ &= \begin{cases} 1 & 0 < u < 1 \\ 0 & \text{otherwise} \end{cases} \times \begin{cases} \frac{1}{u} & 0 < v < u \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} \frac{1}{u} & 0 < v < u < 1 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

As Required

(b) Find the marginal PDF of V

Solution Note that $v < u < 1$ (so the lower bound is v)

$$\begin{aligned} f_V(v) &= \int_v^1 \frac{1}{u} du \\ &= \log(u)|_v^1 \\ &= \log(1) - \log(v) \\ &= \log(\frac{1}{v}) \end{aligned}$$

As Required

Honours Questions

7. Suppose $U \sim U(0, 1)$ and $V|u \sim U(0, u)$

(a) State $E(U)$ and $var(U)$

Solution

$$E(U) = \frac{1 + 0}{2} = \frac{1}{2}, \quad var(U) = \frac{(1 - 0)^2}{12} = \frac{1}{12}$$

As Required

(b) Find $E(V)$

Solution using 6b

$$\begin{aligned} E(V) &= \int_V v \log(\frac{1}{v}) dv \\ &= \int_0^1 v(-\log(v)) dv \\ &\text{use integration by parts: } f = \log(v) \quad g' = v \\ &= -\frac{v^2 \log(v)}{2}|_0^1 + \int_0^1 \frac{v}{2} dv \\ &= \frac{-\log(1)}{2} + \frac{1}{4} = \frac{1}{4} \end{aligned}$$

As Required

(c) Find $var(V)$

Solution

$$\begin{aligned}
var(V) &= E\left((v - E(v))^2\right) \\
&= E\left((v - \frac{1}{4})^2\right) \\
&= E(v^2 - \frac{1}{2}v + \frac{1}{16}) \\
&= E(v^2) - \frac{1}{2}E(v) + \frac{1}{16} \\
&= E(v^2) - \frac{1}{16} \\
E(v^2) &= \int_V v^2 \log(\frac{1}{v}) dv = \frac{1}{9} \text{ using matlab} \\
\implies var(V) &= \frac{1}{9} - \frac{1}{16} = \frac{7}{144}
\end{aligned}$$

As Required

(d) Find $cor(U,V)$

Solution Recall $cor(U,V) = \frac{cov(U,V)}{var(U)var(V)}$ **As Required**

(e) Use R to simulate $1,000,000$ pairs of observations of (U,V) and use your simulations to demonstrate the marginal distribution of V from question 6 and the moment calculations in this question.

Solution The code used:

```
n=1000000
UV=matrix(data=NA,nrow=n,ncol=2)
for( i in 1:n){
  u = runif(1)
  v = runif(1,max=u)
  UV[i,1]=u
  UV[i,2]=v
}
```

As Required