Assignment 1,2

Problem 53, 58 of UGC Math 2019

Vishnu G

Indian Institute of Technology Hyderabad

Table of contents

Problem 53

Suppose (X,Y) follows bivariate normal distribution with means $\mu 1 \mu 2$, standard deviations $\sigma 1$,$\sigma 2$ and correlation coefficient ρ, where all parameters are un-known. Then, testing Ho: $\sigma 1 = \sigma 2$ is equivalent to testing the independence of
1.) X and Y
2.) X and X-Y
3.) X+Y and Y
4.) X+Y and X-Y

Answer: 4

Bivariate Normal Distribution:

Distribution XY in 2 dimensions is said to be Bi-variate normal distribution if X and Y are independently normal distributions and are jointly normal in 2 dimensions.

Correlation of of X,Y $\rho[X, Y] = \frac{\sigma_{[X,Y]}}{\sqrt{\sigma_X^2 \times \sigma_Y^2}}$

If X,Y are independent then $\sigma_{XY} = 0$

So, for testing the independence, if we assume they are independent and then try to prove if $\sigma_{XY} = 0$, for all the options, we can find out the answer.

Known equations

$$\sigma_{[X+Y,Z]} = \sigma_{[X,Z]} + \sigma_{[Y,Z]} \tag{1}$$

$$\sigma_{[X,X]} = \sigma^2_{[X]} \tag{2}$$

$$if \quad \sigma[X,Y] = \sigma[X]^2 \quad then \quad X = Y \tag{3}$$

$$\sigma_{[X,Y]} = 0 \implies X, Y are\ dependent \tag{4}$$

(Proofs for above equations are given in the end)

Testing Independence of X,Y

1. If $\sigma_{[X,Y]}$ is equal to 0, from Eq 4 it means that $Y = X$, which means both of them are dependent

$$\begin{aligned}\sigma[X,Y] &= E((X-\mu_X)(Y-\mu_Y)) \\ &= E[(XY) - \mu_X Y - \mu_Y X + \mu_X\mu_Y] \\ &= E(XY) - \mu_X E(Y) - \mu_Y E(X) + \mu_X\mu_y \\ &= E(XY) - \mu_X\mu_Y - \mu_Y\mu_X + \mu_X\mu_Y\end{aligned}$$

$$\sigma[X,Y] = 0 \implies E(XY) = E(X)E(Y)$$

if $E(XY) = E(X)E(Y)$ *then* $\rho(X,Y) = 1 \implies X, Y$ are dependant

Testing Independence of X and X-Y

1. $$\begin{aligned} \sigma_{[X-Y,X]} &= \sigma_{[X,X]} - \sigma_{[X,Y]} \\ &= \sigma_X^2 - \sigma_{[X,Y]} \end{aligned}$$

2. if $\sigma_X^2 = \sigma_{[X,Y]}$ then it means $Y = X$ From Eq 3 which means they are dependant

Testing independence of X+Y, Y

1.

$$\begin{aligned}\sigma_{[X+Y,X]} &= \sigma_{[X,X]} + \sigma_{[X,Y]} \\ &= \sigma_X^2 + \sigma_{[X,Y]}\end{aligned}$$

2. if $\sigma_X^2 = \sigma_{[X,Y]}$ then it means $Y = -X$ which means they are dependant

Testing for Independence of X+Y and X-Y

1.

$$
\begin{aligned}
&\sigma_{[X+Y,X-Y]} \\
&= \sigma_{[(X+Y),X]} - \sigma_{[(X+Y),Y]} \\
&= \sigma_{[X,X]} + \sigma_{[X,Y]} - \sigma_{[X,X]} - \sigma_{[X,Y]} \\
&= \sigma_X^2 - \sigma_Y^2
\end{aligned}
$$

2. Now testing for $\sigma_X = \sigma_Y \implies \sigma_{[X+Y,X-Y]} = 0$
3. $\Rightarrow \rho[X+Y, X-Y]$ = 0
4. Hence testing for $\sigma_X = \sigma_Y \implies X+Y, X-Y$ are independent.

$\sigma_{[X+Y,Z]} = \sigma_{[X,Z]} + \sigma_{[Y,Z]}$

$$\begin{aligned}
\sigma_{[X+Y,Z]} &= E((X + Y - \mu_{X+Y}) \times (Z - \mu_Z)) \\
&= E((X + Y - \mu_X - \mu_Y) \times (Z - \mu_Z)) \\
&= E(XZ - X\mu_z + YZ - Y\mu_Z - Z\mu_X + \mu_X\mu_Z - \mu_Y Z + \mu_Y\mu_Z) \\
&= E((XZ - \mu_Z X - \mu_X Z + \mu_z\mu_x) + (YZ - Y\mu_Z - \mu_Y Z + \mu_Y\mu_Z)) \\
&= E((X - \mu_X)(Z - \mu_Z) + (Y - \mu_Y)(Z - \mu_Z)) \\
&= E((X - \mu_X)(Z - \mu_Z)) + E((Y - \mu_Y)(Z - \mu_Z)) \\
&= \sigma_{[X,Z]} + \sigma_{[Y,Z]}
\end{aligned}$$

$\sigma_{[X,Y]} = \sigma^2_{[X]}$

$$
\begin{aligned}
\sigma_{[X,Y]} &= E[(X - \mu_X)(Y - \mu_Y)], \quad if \quad Y = X \\
\sigma_{[X,X]} &= E[(X - \mu_X)(X - \mu_X)] \\
&= E[(X - \mu_X)^2] \\
&= \sigma^2_X
\end{aligned}
$$

Assignment 2

Question

A sample of size $n = 2$ is drawn from a population of size $N = 4$ using probability proportional to size without replacement sampling scheme, where the probabilities proportional to size are
The probability of inclusion of unit 1 in the sample is

i	1	2	3	4
Pi	0.4	0.2	0.2	0.2

1. 0.4
2. 0.6
3. 0.7
4. 0.75

Answer : Option - 3 - (0.7)

Definition of PPS without replacement

1. In a simple random sampling scheme, every unit in the population, has an equal probability of getting selected for a sample subset.
2. But when if in a given population sizes of different classes are different, then probability that a sample belonging to a certain class (or unit) differs.
3. Immediate repetition of a sample from same class is not allowed in PPSWOR

Solution

For finding the inclusion probability of unit 1 in a sample of 2, we need find in how many ways unit 1 can be included in a sample of 2.

1. U_1 is included in 1^st draw, and not in second draw
2. U_1 is not included in 1^st draw and is included in second draw.

Probability of U_1 being included in 2^{nd} draw

$P_{i(2)}$ which is probability that i^{th} unit is drawn from the population in the second draw. It can be derived as follows:
$P_{i(2)}$ can occur in following possible ways:

- U_1 is selected in 1^{st} draw and S_i is selected at 2 draw
- U_2 is selected in 1^{st} draw and S_i is selected at 2 draw
- .
- .
- U_{i-1} is selected in 1^{st} draw and S_i is selected at 2 draw
- U_{i-1} is selected in 1^{st} draw and S_i is selected at 2 draw
- U_N is selected in 1^{st} draw and S_i is selected at 2 draw

Probability of U_1 being included in 2^{nd} draw

It has to be noted that S_i is being skipped in for first draw because repetition is not allowed in PPS without replacement scheme.
S_1 is selected in 1^{st} draw and S_i is selected at 2 draw is derived as
Let S_1 is selected in 1^{st} draw be Event A, S_i is selected at 2 as Event B

$$P(A, B) = P(A) \times P(B/A), \quad where, P(A) = P_1,$$
$$P(B/A) = \frac{P_i}{1 - (P_1)}$$

P(B/A) = Population of P_i over total remaining population, which is $1 - P_1$, because P_1 can not be selected again

Probability of U_1 being included in 2^{nd} draw

$$
\begin{aligned}
P_{i(2)} &= P1 \times \frac{P_i}{1-P_1} + P2 \times \frac{P_i}{1-P_2} + \dots \\
&\quad + P_{i-1} \times \frac{P_i}{1-P_{i-1}} + P_{i+1} \times \frac{P_i}{1-P_{i+1}} + P_N \times \frac{P_i}{1-P_N} \\
P_{i(2} &= \sum_{j=i}^{N} P_j \times \frac{P_i}{1-P_j} - P_i \times \frac{P_i}{1-P_i}
\end{aligned}
$$

$$
\begin{aligned}
P_{i(2)} &= \sum_{j=i}^{N} P_j \times \frac{P_i}{1-P_j} - P_i \times \frac{P_i}{1-P_i} \\
&= 0.4 \times \left(\frac{0.2}{1-0.2} + \frac{0.2}{1-0.2} + \frac{0.2}{1-0.2} + \frac{0.4}{1-0.4} - \frac{0.4}{1-0.4} \right) \\
&= 0.4 \times \left(\frac{0.6}{0.8} \right) \\
&= 0.3
\end{aligned}
$$

Total inclusion probability of unit 1

- The probability of inclusion of unit 1 in the sample is $P_{i(1)} + P_{i(2)}$
- 0.3+0,4 = 0.7