



You:

Manufacturer of product 'A' claims that average life of the product is 120 months with variance 25 months. A random sample of 10 of these is with mean 125 months and standard deviation 4 months. Validate the claim of the manufacturer at 1% level of significance



ChatGPT:

Certainly, let's validate the manufacturer's claim using hypothesis testing. The goal is to determine if the sample data provides enough evidence to reject the manufacturer's claim about the average life of product 'A'.

Step 1: State the Hypotheses

- Null Hypothesis $H_0: \mu = 120$ (The average life of the product is 120 months)
- Alternative Hypothesis $H_1: \mu \neq 120$ (The average life of the product is not 120 months)

Step 2: Choose the Level of Significance

The level of significance α is given as 1% or 0.01.

Step 3: Identify the Test Statistic

Since the sample size $n = 10$ is less than 30 and we know the population variance, we'll use the Z-test for this hypothesis test.

The formula for the Z-test statistic is:

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$Z =$

$\frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$

Where:

- $\bar{x} = 125$ (sample mean)
- $\mu = 120$ (population mean under null hypothesis)
- $\sigma = \sqrt{25} = 5$ (population standard deviation)
- $n = 10$ (sample size)

Step 4: Calculate the Test Statistic

$$Z = \frac{125 - 120}{\frac{5}{\sqrt{10}}}$$

$Z =$

$\frac{125 - 120}{\frac{5}{\sqrt{10}}}$

$Z =$

$\frac{5}{\sqrt{10}}$

$Z =$

≈ 3.16

$$Z = \sqrt{10} \approx 3.16$$

Step 5: Make a Decision

Now, we need to find the critical Z-value for a two-tailed test at $\alpha = 0.01$.

The critical Z-value for $\alpha/2 = 0.005$ in each tail is approximately ± 2.576 .

Since the calculated Z-value 3.16 is greater than the critical Z-value 2.576, we reject the null hypothesis.

Conclusion

At the 1% level of significance, we have enough evidence to reject the manufacturer's claim that the average life of the product is 120 months. The sample data suggests that the average life is different from 120 months.

Does this clarify how to validate the manufacturer's claim using hypothesis testing? Would you like to go over any part in more detail?



You:

It is claimed that player A is better than Player B in IPL matches during IPL bidding for players. During previous IPL season both played 18 matches with average runs scored 85 and 75 with standard deviations 9 and 2 respectively. Validate the claim using appropriate statistical procedure at 1% level of significance



ChatGPT:

To validate the claim that Player A is better than Player B in terms of average runs scored in IPL matches, we can use a hypothesis test for the difference between two means. Since we have sample data for both players, we can use the Z-test for comparing two means.

Step 1: State the Hypotheses

- Null Hypothesis H_0 : $\mu_A = \mu_B$ (The average runs scored by both players are the same)
- Alternative Hypothesis H_1 : $\mu_A > \mu_B$ (Player A scores more runs on average than Player B)

Step 2: Choose the Level of Significance

The level of significance α is given as 1% or 0.01.

Step 3: Identify the Test Statistic

We'll use the Z-test statistic for comparing two means, which is given by:

$$Z = \frac{(\bar{x}_A - \bar{x}_B) - (\mu_A - \mu_B)}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}}$$

$$Z = \frac{85 - 75 - 0}{\sqrt{\frac{9^2}{18} + \frac{2^2}{18}}}$$

$$Z = \frac{(85 - 75) - 0}{\sqrt{\frac{81}{18} + \frac{4}{18}}}$$

Where:

- $\bar{x}_A = 85$ and $\bar{x}_B = 75$ are the sample means for Player A and Player B, respectively.
- $\mu_A - \mu_B = 0$ under the null hypothesis.
- $s_A = 9$ and $s_B = 2$ are the sample standard deviations for Player A and Player B, respectively.
- $n_A = n_B = 18$ are the sample sizes for both players.

Step 4: Calculate the Test Statistic

$$Z = \frac{(85 - 75) - 0}{\sqrt{\frac{9^2}{18} + \frac{2^2}{18}}}$$

$$Z = \frac{10}{\sqrt{4.5 + 0.2222}}$$

$$Z = \frac{10}{\sqrt{4.7222}}$$

$$Z = \frac{10}{\sqrt{4.7222}}$$

$$Z = \frac{10}{2.173}$$

$$Z = \frac{10}{2.173}$$

$$Z = \frac{10}{2.173}$$

$$Z = 4.59$$

$$Z = 4.59$$

$$Z = \frac{10}{\sqrt{4.7222}}$$

$$Z = 4.7222$$

$$\sqrt{10}$$

$$Z = \frac{10}{2.174}$$

$$Z = 2.17410$$

$$Z \approx 4.6$$

$$Z \approx 4.6$$

Step 5: Make a Decision

For a one-tailed test at $\alpha = 0.01$, the critical Z-value is approximately 2.332.33.

Since the calculated Z-value 4.64.6 is greater than the critical Z-value 2.332.33, we reject the null hypothesis.

Conclusion

At the 1% level of significance, we have enough evidence to reject the null hypothesis that the average runs scored by both players are the same. The sample data suggests that Player A scores more runs on average than Player B, supporting the claim made during the IPL bidding.

Would you like to discuss any part of this analysis in more detail?



You:

Formulate a suitable hypothesis and validate it by using appropriate statistical procedure based on the following data. (At 5% level of significance) "Before pandemic 200 out of 750 students are attending online classes whereas after pandemic 350 out of 800 are doing the same"



ChatGPT:

To validate whether the pandemic has had a significant impact on the proportion of students attending online classes, we can use a hypothesis test for comparing two proportions.

Step 1: State the Hypotheses

- Null Hypothesis H_0 : $p_{\text{before}} = p_{\text{after}}$ (The proportion of students attending online classes is the same before and after the pandemic)
- Alternative Hypothesis H_1 : $p_{\text{before}} \neq p_{\text{after}}$ (The proportion of students attending online classes is different before and after the pandemic)

Step 2: Choose the Level of Significance

The level of significance α is given as 5% or 0.05.

Step 3: Identify the Test Statistic

We'll use the Z-test statistic for comparing two proportions, which is given by:

$$Z = \frac{\hat{p}_{\text{before}} - \hat{p}_{\text{after}}}{\sqrt{p(1-p)\left(\frac{1}{n_{\text{before}}} + \frac{1}{n_{\text{after}}}\right)}}$$

$$Z = \frac{\hat{p}_{\text{before}} - \hat{p}_{\text{after}}}{\sqrt{p(1-p)\left(\frac{1}{n_{\text{before}}} + \frac{1}{n_{\text{after}}}\right)}}$$

$$\sqrt{\frac{\hat{p}_{\text{before}} - \hat{p}_{\text{after}}}{p(1-p)\left(\frac{1}{n_{\text{before}}} + \frac{1}{n_{\text{after}}}\right)}}$$

Where \hat{p}_{before} and \hat{p}_{after} are the sample proportions, and p is the pooled proportion:

$$p = \frac{x_{\text{before}} + x_{\text{after}}}{n_{\text{before}} + n_{\text{after}}}$$

$$p = \frac{x_{\text{before}} + x_{\text{after}}}{n_{\text{before}} + n_{\text{after}}}$$

Step 4: Calculate the Test Statistic

- $\hat{p}_{\text{before}} = \frac{200}{750} \approx 0.267$

- $\hat{p}_{\text{after}} = \frac{350}{800} \approx 0.438$ $p^{\text{after}} = 800350 \approx 0.438$
- $n_{\text{before}} = 750$ $n_{\text{before}} = 750$
- $n_{\text{after}} = 800$ $n_{\text{after}} = 800$

First, calculate the pooled proportion p :

$$p = \frac{200 + 350}{750 + 800} = \frac{550}{1550} \approx 0.355$$

$$p = 750 + 800200 + 350 = 1550550 \approx 0.355$$

Now, calculate the Z-value:

$$Z = \frac{0.267 - 0.438}{\sqrt{0.355(1 - 0.355)\left(\frac{1}{750} + \frac{1}{800}\right)}}$$

$$Z = 0.355(1 - 0.355)(7501 + 8001)$$

$$\sqrt{0.267 - 0.438}$$

$$Z = \frac{-0.171}{\sqrt{0.355 \times 0.645 \times \left(\frac{1}{750} + \frac{1}{800}\right)}}$$

$$Z = 0.355 \times 0.645 \times (7501 + 8001)$$

$$\sqrt{-0.171}$$

$$Z \approx \frac{-0.171}{\sqrt{0.229 \times (0.00133 + 0.00125)}}$$

$$Z \approx 0.229 \times (0.00133 + 0.00125)$$

$$\sqrt{-0.171}$$

$$Z \approx \frac{-0.171}{\sqrt{0.229 \times 0.00258}}$$

$$Z \approx 0.229 \times 0.00258$$

$$\sqrt{-0.171}$$

$$Z \approx \frac{-0.171}{\sqrt{0.00059082}}$$

$$Z \approx 0.00059082$$

$$\sqrt{-0.171}$$

$$Z \approx \frac{-0.171}{0.0243}$$

$$Z \approx 0.0243 - 0.171$$

$$Z \approx -7.03$$

$$Z \approx -7.03$$

Step 5: Make a Decision

For a two-tailed test at $\alpha = 0.05$, the critical Z-values are approximately ± 1.96 .

Since the calculated Z-value -7.03 is less than the lower critical Z-value -1.96 , we reject the null hypothesis.

Conclusion

At the 5% level of significance, we have enough evidence to reject the null hypothesis that the proportion of students attending online classes is the same before and after the pandemic. The sample data suggests that the proportion has changed significantly, indicating that more students are attending online classes after the pandemic.

Would you like to discuss any part of this analysis in more detail?



You:

Consider the following data. Use an appropriate statistical understanding (numerical) to take the decision to proceed further or not to build a simple linear regression ($y = f(x)$) model for prediction. If yes, use it predict y when $x = 10$. X 6 4 9 5 3 Y 8 7 6 9 6



To decide whether to proceed with building a simple linear regression model, we can first calculate the correlation coefficient r to measure the strength and direction of the relationship between X and Y .

Step 1: Calculate the Correlation Coefficient r

The formula for r is:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

First, let's find the sums needed for the formula:

- $\sum x = 6 + 4 + 9 + 5 + 3 = 27$
- $\sum y = 8 + 7 + 6 + 9 + 6 = 36$
- $\sum xy = 6 \times 8 + 4 \times 7 + 9 \times 6 + 5 \times 9 + 3 \times 6 = 48 + 28 + 54 + 45 + 18 = 193$
- $\sum x^2 = 6^2 + 4^2 + 9^2 + 5^2 + 3^2 = 36 + 16 + 81 + 25 + 9 = 167$
- $\sum y^2 = 8^2 + 7^2 + 6^2 + 9^2 + 6^2 = 64 + 49 + 36 + 81 + 36 = 266$
- $n = 5$ (number of data points)

Now, we can plug these into the formula:

$$r = \frac{5 \times 193 - 27 \times 36}{\sqrt{(5 \times 167 - 27^2)(5 \times 266 - 36^2)}}$$

$$r = \frac{5 \times 193 - 27 \times 36}{\sqrt{(5 \times 167 - 27^2)(5 \times 266 - 36^2)}}$$

$$r = \frac{5 \times 193 - 27 \times 36}{\sqrt{(5 \times 167 - 27^2)(5 \times 266 - 36^2)}}$$

$$r = \frac{965 - 972}{\sqrt{(835 - 729)(1330 - 1296)}}$$

$$r = \frac{965 - 972}{\sqrt{(835 - 729)(1330 - 1296)}}$$

$$r = \frac{965 - 972}{\sqrt{(835 - 729)(1330 - 1296)}}$$

$$r = \frac{-7}{\sqrt{(106)(34)}}$$

$$r = \frac{-7}{\sqrt{(106)(34)}}$$

$$r = \frac{-7}{\sqrt{(106)(34)}}$$

$$r = \frac{-7}{\sqrt{3604}}$$

$$r = \frac{-7}{\sqrt{3604}}$$

$$r = \frac{-7}{\sqrt{3604}}$$

$$r = \frac{-7}{60.03}$$

$$r = \frac{-7}{60.03}$$

$$r \approx -0.117$$

The correlation coefficient r is approximately -0.117, which indicates a very weak negative correlation between X and Y .

Decision

Given the weak correlation, it may not be advisable to proceed with building a simple linear regression model for prediction, as the model is unlikely to provide accurate predictions.

However, if you still want to proceed for educational purposes, we can calculate the linear regression equation $y = mx + b$, where m is the slope and b is the y-intercept.

Would you like to proceed with that?



yes



Alright, let's proceed with building the simple linear regression model for educational purposes.

Step 2: Calculate the Slope m and Y-Intercept b

The formula for the slope m is:

$$m = \frac{n(\sum xy) - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2}$$

$$m = \frac{n\sum x^2 - (\sum x)^2}{n(\sum xy) - (\sum x)(\sum y)}$$

We already have the sums from the correlation calculation, so we can plug them into the formula:

$$m = \frac{5 \times 193 - 27 \times 36}{5 \times 167 - 27^2}$$

$$m = \frac{5 \times 193 - 27 \times 36}{5 \times 167 - 27^2}$$

$$m = \frac{965 - 972}{835 - 729}$$

$$m = \frac{965 - 972}{835 - 729}$$

$$m = \frac{-7}{106}$$

$$m = \frac{-7}{106}$$

$$m \approx -0.066$$

$$m \approx -0.066$$

The formula for the y-intercept b is:

$$b = \frac{\sum y - m(\sum x)}{n}$$

$$b = \frac{\sum y - m(\sum x)}{n}$$

$$b = \frac{36 - (-0.066 \times 27)}{5}$$

$$b = \frac{36 - (-0.066 \times 27)}{5}$$

$$b = \frac{36 + 1.782}{5}$$

$$b = \frac{36 + 1.782}{5}$$

$$b = \frac{37.782}{5}$$

$$b = \frac{37.782}{5}$$

$$b \approx 7.556$$

$$b \approx 7.556$$

Step 3: Formulate the Linear Regression Equation

The linear regression equation $y = mx + b$ becomes:

$$y = -0.066x + 7.556$$

$$y = -0.066x + 7.556$$

Step 4: Predict y when $x = 10$

To predict y when $x = 10$, we can plug x into the equation:

$$y = -0.066 \times 10 + 7.556$$

$$y = -0.066 \times 10 + 7.556$$

$$y = -0.66 + 7.556$$

$$y = -0.66 + 7.556$$

$$y = 6.896$$

$$y = 6.896$$

So, according to this simple linear regression model, the predicted value of y when $x = 10$ is approximately 6.896.

Keep in mind that given the weak correlation, this prediction is likely not very accurate. However, it serves as a good exercise in understanding how to build a simple linear regression model.

Would you like to discuss any part of this analysis in more detail?



You:

Try to model the following time series data by using two models mentioned. Suggest the better of these two to be used for forecast. a) Model 1: Simple moving averages model with $k = 3$ b) Model 2: Weighted moving averages with weights 0.3, 0.2 and 0.1 (i.e latest will be given more weightage and so on)

Month 1 2 3 4 5 6 7 Sales(in lakhs) 10 12 11 6 8 9 11 Number of questions:4 Number of Pages: 2



ChatGPT:

To compare the two models, we'll first calculate the forecasts for the given time series data using both the Simple Moving Averages model and the Weighted Moving Averages model. Then we'll assess which model seems to be better for forecasting.

Data

The sales data in lakhs for 7 months is: [10, 12, 11, 6, 8, 9, 11]

Model 1: Simple Moving Averages with $k = 3$

For this model, we take the average of the last 3 data points to forecast the next point. The forecast starts from the 4th data point.

- Forecast for Month 4: $\frac{10 + 12 + 11}{3} = 11$
- Forecast for Month 5: $\frac{12 + 11 + 6}{3} = 9.67$
- Forecast for Month 6: $\frac{11 + 6 + 8}{3} = 8.33$
- Forecast for Month 7: $\frac{6 + 8 + 9}{3} = 7.67$

Model 2: Weighted Moving Averages with weights [0.3, 0.2, 0.1]

For this model, we multiply each of the last 3 data points by their respective weights and sum them up to forecast the next point. The forecast starts from the 4th data point.

- Forecast for Month 4: $10 \times 0.1 + 12 \times 0.2 + 11 \times 0.3 = 1 + 2.4 + 3.3 = 6.7$
- Forecast for Month 5: $12 \times 0.1 + 11 \times 0.2 + 6 \times 0.3 = 1.2 + 2.2 + 1.8 = 5.2$
- Forecast for Month 6: $11 \times 0.1 + 6 \times 0.2 + 8 \times 0.3 = 1.1 + 1.2 + 2.4 = 4.7$
- Forecast for Month 7: $6 \times 0.1 + 8 \times 0.2 + 9 \times 0.3 = 0.6 + 1.6 + 2.7 = 4.9$

Model Assessment

To assess the models, we can calculate the Mean Absolute Error (MAE) for each model. The MAE is the average of the absolute differences between the observed actual outcomes and the forecasts.

- MAE for Model 1: $\frac{|11 - 6| + |9.67 - 8| + |8.33 - 9| + |7.67 - 11|}{4} = \frac{5 + 1.67 + 0.67 + 3.33}{4} = \frac{10.67}{4} = 2.67$
- MAE for Model 2: $\frac{|6.7 - 6| + |5.2 - 8| + |4.7 - 9| + |4.9 - 11|}{4} = \frac{0.7 + 2.8 + 4.3 + 6.1}{4} = \frac{13.9}{4} = 3.475$

Conclusion

Model 1 (Simple Moving Averages with $k = 3$) has a lower MAE (2.67) compared to Model 2 (Weighted Moving Averages) with an MAE of 3.475. Therefore, Model 1 would be a better choice for forecasting future sales based on the given data.

Would you like to discuss any part of this analysis in more detail?



You:

A random variable X follows binomial distribution with $n = 1000$ and $p = 0.001$. Then find i) $P(X > 700)$ ii) $P(X < 300)$ iii) $P(300 < X < 700)$.



ChatGPT:

In a binomial distribution with $n = 1000$ and $p = 0.001$, the parameters are quite extreme. The expected value $\mu = np = 1000 \times 0.001 = 1$ and the variance $\sigma^2 = np(1-p) = 1000 \times 0.001 \times 0.999 = 0.999$. The standard deviation $\sigma = \sqrt{0.999} \approx 1$.

Given these parameters, the probabilities you're asking for are essentially zero for all practical purposes. Here's why:

i) $P(X > 700)$

The probability of getting more than 700 successes in 1000 trials with a success rate of 0.001 is practically zero. In real-world terms, this would be an extremely unlikely event given the very low probability of success on each trial.

ii) $P(X < 300)$

Given that the expected number of successes is 1, the probability of getting fewer than 300 successes is essentially 1. Again, in real-world terms, it's almost certain that you'll get fewer than 300 successes with these parameters.

iii) $P(300 < X < 700)$

Similar to the first case, the probability of getting between 300 and 700 successes with these parameters is practically zero.

In summary, the probabilities for these events are so extreme that they are essentially zero or one, given the parameters of the binomial distribution in this case. Normally, you could use the binomial probability formula or a normal approximation for large n , but in this case, the extreme values make those calculations unnecessary.

Would you like to discuss this further?



You:

Consider the following probabilities. $x_1 \ x_2 \ y \ P(x_1, x_2, y)$

1	1	0	0.20
1	1	1	0.15
1	0	1	0.02
0	0	0	0.15
0	0	1	0.25
1	0	0	0.10
0	1	0	0.05
0	1	1	0.08

Can you first format these values into a proper table? Then, if possible find the following. If not, validate your decision i) $P(x_1)$ ii) $P(x_2)$ iii) $P((x_1, x_2) / y = 0)$ iv) $P((x_1, x_2) / y = 1)$ v) $P(x_1 = 0 / x_2 = 1)$



ChatGPT:

Certainly, let's first organize the given probabilities into a table for better visualization:

x_1	x_2	y	$P(x_1, x_2, y)$
1	1	0	0.20
1	1	1	0.15
1	0	1	0.02
0	0	0	0.15
0	0	1	0.25
1	0	0	0.10
0	1	0	0.05
0	1	1	0.08

Before proceeding, let's verify that the probabilities sum to 1:

$$0.20 + 0.15 + 0.02 + 0.15 + 0.25 + 0.10 + 0.05 + 0.08 = 1.00$$

Great, the probabilities sum to 1, so we can proceed with the calculations.

i) $P(x_1)$

To find $P(x_1)$, we sum the probabilities where $x_1 = 1$ and $x_1 = 0$ separately:

- $P(x_1 = 1) = 0.20 + 0.15 + 0.02 + 0.10 = 0.47$
- $P(x_1 = 0) = 0.15 + 0.25 + 0.05 + 0.08 = 0.53$

ii) $P(x_2)$

Similarly, for $P(x_2)$:

- $P(x_2 = 1) = 0.15 + 0.02 + 0.05 + 0.08 = 0.28$
- $P(x_2 = 0) = 0.20 + 0.25 + 0.10 + 0.08 = 0.63$

$$\bullet P(x_2 = 0) = 0.02 + 0.15 + 0.25 + 0.10 = 0.52 \quad P(x_2 = 0) = 0.02 + 0.15 + 0.25 + 0.10 = 0.52$$

iii) $P((x_1, x_2)|y = 0)P((\mathbf{x1}, \mathbf{x2})|y = 0)$

To find this conditional probability, we first find $P(y = 0)P(y = 0)$:

$$P(y = 0) = 0.20 + 0.15 + 0.10 + 0.05 = 0.50 \quad P(y = 0) = 0.20 + 0.15 + 0.10 + 0.05 = 0.50$$

Then, we can find $P((x_1, x_2)|y = 0)P((\mathbf{x1}, \mathbf{x2})|y = 0)$ for each combination of x_1 and x_2 :

$$\begin{aligned} \bullet P((x_1 = 1, x_2 = 1)|y = 0) &= \frac{0.20}{0.50} = 0.4 \quad P((x_1 = 1, x_2 = 1)|y = 0) = 0.50 \cdot 0.20 = 0.10 \\ \bullet P((x_1 = 1, x_2 = 0)|y = 0) &= \frac{0.10}{0.50} = 0.2 \quad P((x_1 = 1, x_2 = 0)|y = 0) = 0.50 \cdot 0.10 = 0.05 \\ \bullet P((x_1 = 0, x_2 = 1)|y = 0) &= \frac{0.05}{0.50} = 0.1 \quad P((x_1 = 0, x_2 = 1)|y = 0) = 0.50 \cdot 0.05 = 0.025 \\ \bullet P((x_1 = 0, x_2 = 0)|y = 0) &= \frac{0.15}{0.50} = 0.3 \quad P((x_1 = 0, x_2 = 0)|y = 0) = 0.50 \cdot 0.15 = 0.075 \end{aligned}$$

iv) $P((x_1, x_2)|y = 1)P((\mathbf{x1}, \mathbf{x2})|y = 1)$

First, find $P(y = 1)P(y = 1)$:

$$P(y = 1) = 0.15 + 0.02 + 0.25 + 0.08 = 0.50 \quad P(y = 1) = 0.15 + 0.02 + 0.25 + 0.08 = 0.50$$

Then, for each combination of x_1 and x_2 :

$$\begin{aligned} \bullet P((x_1 = 1, x_2 = 1)|y = 1) &= \frac{0.15}{0.50} = 0.3 \quad P((x_1 = 1, x_2 = 1)|y = 1) = 0.50 \cdot 0.15 = 0.075 \\ \bullet P((x_1 = 1, x_2 = 0)|y = 1) &= \frac{0.02}{0.50} = 0.04 \quad P((x_1 = 1, x_2 = 0)|y = 1) = 0.50 \cdot 0.02 = 0.01 \\ \bullet P((x_1 = 0, x_2 = 1)|y = 1) &= \frac{0.08}{0.50} = 0.16 \quad P((x_1 = 0, x_2 = 1)|y = 1) = 0.50 \cdot 0.08 = 0.04 \\ \bullet P((x_1 = 0, x_2 = 0)|y = 1) &= \frac{0.25}{0.50} = 0.5 \quad P((x_1 = 0, x_2 = 0)|y = 1) = 0.50 \cdot 0.25 = 0.125 \end{aligned}$$

v) $P(x_1 = 0|x_2 = 1)P(\mathbf{x1} = 0|\mathbf{x2} = 1)$

To find this conditional probability, we first find $P(x_2 = 1)P(x_2 = 1)$, which is 0.48 as calculated earlier.

$$\text{Then, } P(x_1 = 0|x_2 = 1) = \frac{P(x_1 = 0, x_2 = 1)}{P(x_2 = 1)} P(x_1 = 0|x_2 = 1) = P(x_2 = 1)P(x_1 = 0, x_2 = 1)$$

$$P(x_1 = 0|x_2 = 1) = \frac{0.05 + 0.08}{0.48} = \frac{0.13}{0.48} \approx 0.271 \quad P(x_1 = 0|x_2 = 1) = 0.48 \cdot 0.271 = 0.13$$

Would you like to discuss any part of these calculations in more detail?



No



Alright, if you have any more questions in the future, feel free to ask. Have a great day!
