Statistical Significance

Dr. Haydar Raheem Hmoud

What is Hypothesis?

A hypothesis is an assumption that is made based on some evidence. This is the initial point of any investigation that translates the research questions into predictions. It includes components like variables, population and the relation between the variables. A research hypothesis is a hypothesis that is used to test the relationship between two or more variables.

Characteristics of Hypothesis

Following are the characteristics of the hypothesis:

- The hypothesis should be clear and precise to consider it to be reliable.
- If the hypothesis is a relational hypothesis, then it should be stating the relationship between variables.
- The hypothesis must be specific and should have scope for conducting more tests.
- The way of explanation of the hypothesis must be very simple and it should also be understood that the simplicity of the hypothesis is not related to its significance.

Sources of Hypothesis

Following are the sources of hypothesis:

- The resemblance between the phenomenon.
- Observations from past studies, present-day experiences and from the competitors.
- Scientific theories.
- General patterns that influence the thinking process of people.

Types of Hypothesis

There are six forms of hypothesis and they are:

- Simple hypothesis
- Complex hypothesis
- Directional hypothesis
- Non-directional hypothesis
- **Null hypothesis**
- Associative and casual hypothesis

Null Hypothesis

It provides a statement which is contrary to the hypothesis. It's a negative statement, and there is no relationship between independent and dependent variables. The symbol is denoted by "H_0".

Example: In a clinical trial of a new anti-hypertensive drug, the null hypothesis would state that there is no difference in effect when comparing the new drug to the current standard treatment.

Functions of Hypothesis

Following are the functions performed by the hypothesis:

- Hypothesis helps in making an observation and experiments possible.
- It becomes the start point for the investigation.
- Hypothesis helps in verifying the observations.
- It helps in directing the inquiries in the right direction.

Hypothesis Testing

Investigators conducting studies need research questions and hypotheses to guide analyses. Starting with broad research questions (RQs), investigators then identify a gap in current clinical practice or research. Any research problem or statement is grounded in a better understanding of relationships between two or more variables.

Example of Research Question: Is Drug 23 an effective treatment for Disease A?

Research questions **do not directly** imply specific guesses or predictions; we must formulate research hypotheses. A hypothesis is a predetermined declaration regarding the research question in which the investigator(s) makes a precise, educated guess about a study outcome. This is sometimes called the **alternative hypothesis** and ultimately allows the researcher to take a stance based on experience or insight from medical literature. An example of a hypothesis is below.

Research Hypothesis: Drug 23 will significantly reduce symptoms associated with Disease A compared to Drug 22.

The null hypothesis states that there is **no statistical difference** between groups based on the stated research hypothesis.

P Values

P values are used in research to determine whether the sample estimate is significantly different from a hypothesized value. The p-value is the probability that the observed effect within the study would have occurred by **chance** if, in reality, there was no true effect. Conventionally, data yielding a $p<0.05$ or $p<0.01$ is considered statistically **significant**. While some have debated that the 0.05 level should be lowered, it is still universally practiced. Hypothesis testing allows us to determine the size of the effect.

For example: Individuals who were prescribed Drug 23 experienced fewer symptoms ($M = 1.3$, $SD = 0.7$) compared to individuals who were prescribed Drug 22 ($M = 5.3$, $SD = 1.9$). This finding was statistically significant, $p= 0.02$.

If the threshold had been set at 0.05, the null hypothesis (that there was no relationship) should be rejected, and we should conclude significant differences. Noticeably, as can be seen in the two statements above, some researchers will report findings with < or > and others will provide an exact p-value (0.000001) but never zero. When examining research, readers should understand how p values are reported. The best practice is to report all p values for all variables within a study design, rather than only providing p values for variables with significant findings. The inclusion of all p values provides evidence for study validity and limits suspicion for selective reporting/data mining.

To test a hypothesis, researchers obtain data on a representative sample to determine whether to reject or fail to reject a null hypothesis. In most research studies, it is not feasible to obtain data for an entire population. Using a sampling procedure allows for statistical inference, though this involves a certain possibility of error.

Chi-Square Test Definition

A chi-square test is a statistical test that is used to compare **observed** and **expected** results. The goal of this test is to identify whether a disparity between actual and predicted data is due to chance or to a **link** between the variables under consideration. As a result, the chi-square test is an ideal choice for aiding in our understanding and interpretation of the connection between our two **categorical variables**.

A chi-square test or comparable **nonparametric** test is required to test a hypothesis regarding the distribution of a categorical variable. Categorical variables, which indicate categories such as Blood groups or Gender, can be **nominal** or **ordinal**. They cannot have a normal distribution since they can only have a few particular values.

It is used to calculate the difference between two categorical variables, which are:

- As a result of chance or
- Because of the relationship

Formula For Chi-Square Test

$$\chi_c^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

Where

c = Degrees of freedom

O = Observed Value

E = Expected Value

The degrees of freedom in a statistical calculation represent the number of variables that can vary in a calculation. The degrees of freedom can be calculated to ensure that chi-square tests are statistically valid.

Example : Consider a data sample consisting of five positive integers. The values of the five integers must have an average of six. If four of the items within the data set are {3, 8, 5, and 4}, the fifth number must be 10. Because the first four numbers can be chosen at random, the degrees of freedom is four.

The Observed values are those you gather yourselves.

The expected values are the frequencies expected, based on the null hypothesis.

Example: A researcher conducts a survey to investigate the severity of Covid-19 patients in residents of different cities in Iraq. Separate random samples of residents were evaluated as shown in the two-way table which summarizes the condition of these patients. Do these data provide convincing evidence at the α = 0.05 level that the distributions of severity of symptoms differ for residents of Baghdad, Mosul, and Basrah?

	Baghdad	Mosul	Basrah	Total
Not sever	690	1177	702	2569
Sever	250	242	221	713
Very Sever	40	28	50	118
Life threatening	20	13	30	63
Total	1000	1460	1003	3463

1) STATE the Hypothesis

H_0: There is **no difference** in the true distributions of severity of symptoms for residents

H_a: There is **a difference** in the true distributions of severity of symptoms for residents

2) Calculate Expected values

Calculating Expected Counts for a Chi-Square Test Based on Data in a Two-Way Table

When H_0 is true, the expected count in any cell of a two-way table is:

$$\textbf{Expected count} = \frac{\text{row total} \cdot \text{column total}}{\text{table total}}$$

=(2569*1000)/3463
=741.8

Observed	Baghdad	Mosul	Basrah	Total
Not sever	690	1177	702	2569
Sever	250	242	221	713
Very Sever	40	28	50	118
Life threatening	20	13	30	63
Total	1000	1460	1003	3463

Expected	Baghdad	Mosul	Basrah	Total
Not sever	741.8	1083.1	744.1	2569
Sever	205.9	300.6	206.5	713
Very Sever	34.1	49.7	34.2	118
Life threatening	18.2	26.6	18.2	63
Total	1000	1460	1003	3463

3) Compute the chi-square test statistic

$$\chi^2 = \sum \frac{(\text{Observed count} - \text{Expected count})^2}{\text{Expected count}}$$

where the sum is over all cells (not including totals) in the two-way table.

$$\chi^2 = \frac{(690 - 741.8)^2}{741.8} + \frac{(1177 - 1083.1)^2}{1083.1} + \cdots = 68.57$$

Observed	Baghdad	Mosul	Basrah	Total
Not sever	690	1177	702	2569
Sever	250	242	221	713
Very Sever	40	28	50	118
Life threatening	20	13	30	63

Expected	Baghdad	Mosul	Basrah	Total
Not sever	741.8	1083.1	744.1	2569
Sever	205.9	300.6	206.5	713
Very Sever	34.1	49.7	34.2	118
Life threatening	18.2	26.6	18.2	63

Statistics	Baghdad	Mosul	Basrah
Not sever	3.6	8.1	2.4
Sever	9.4	11.4	1.0
Very Sever	1.0	9.5	7.3
Life threatening	0.2	6.9	7.6

4) The *P*-value is the area to the right of χ^2 under the chi-square density curve with degrees of freedom
= (num. of rows – 1)(num. of columns – 1).

df = (4 – 1)(3 – 1) = 6

χ^2 **= 68.57**

Using Table of critical values: P-value < 0.001

CONCLUDE

Because the P-value < α = 0.05, we reject H_o. There is convincing evidence that there is a difference in the true distributions of severity for residents of the cities.

Critical values of the Chi-square distribution with *d* degrees of freedom

	Probability of exceeding the critical value						
d	0.05	0.01	0.001	*d*	0.05	0.01	0.001
1	3.841	6.635	10.828	11	19.675	24.725	31.264
2	5.991	9.210	13.816	12	21.026	26.217	32.910
3	7.815	11.345	16.266	13	22.362	27.688	34.528
4	9.488	13.277	18.467	14	23.685	29.141	36.123
5	11.070	15.086	20.515	15	24.996	30.578	37.697
6	12.592	16.812	22.458	16	26.296	32.000	39.252
7	14.067	18.475	24.322	17	27.587	33.409	40.790
8	15.507	20.090	26.125	18	28.869	34.805	42.312
9	16.919	21.666	27.877	19	30.144	36.191	43.820
10	18.307	23.209	29.588	20	31.410	37.566	45.315

INTRODUCTION TO POPULATION GENETICS, Table D.1

common uses:

1) Goodness-of-Fit Test (Chi-squared goodness-of-fit test): This test is used when you want to determine if the observed data fits a specific theoretical or expected distribution. For example, you might want to check if the distribution of blood types in a population follows the expected proportions (e.g., ABO blood group distribution).

Example: Blood Type Distribution

Imagine a researcher wants to determine if the distribution of blood types in a population follows the expected proportions based on the ABO blood group system. According to the ABO blood group system, there are four major blood types: A, B, AB, and O. The expected distribution in a specific population, based on historical data or known genetic frequencies, is as follows:

Blood Type A: 40%

Blood Type B: 10%

Blood Type AB: 4%

Blood Type O: 46%

The researcher collects data from a sample of 500 individuals in the population and records their blood types. The observed distribution in the sample is as follows:

Blood Type A: 200 individuals
Blood Type B: 80 individuals
Blood Type AB: 10 individuals
Blood Type O: 210 individuals

The researcher wants to know if the observed distribution significantly differs from the expected distribution based on the ABO blood group system. To do this, they perform a Chi-squared goodness-of-fit test:

Step 1: Formulate Hypotheses
Null Hypothesis (H0): The observed distribution of blood types in the population follows the expected distribution based on the ABO blood group system.
Alternative Hypothesis (Ha): The observed distribution of blood types in the population does not follow the expected distribution based on the ABO blood group system.
Step 2: Set Significance Level
The researcher decides on a significance level (e.g., $\alpha = 0.05$), which represents the probability of making a Type I error (rejecting the null hypothesis when it's true).

Observed Counts

Blood Type	A	B	AB	O
Observed	200	80	10	210

Step 3: Create a Contingency Table

Blood Type	Expected Count (E)	Observed Count (O)
A	0.40 * 500 = 200	200
B	0.10 * 500 = 50	80
AB	0.04 * 500 = 20	10
O	0.46 * 500 = 230	210

Step 4: Calculate the Chi-squared Statistic

The formula for the Chi-squared statistic is:

$\chi^2 = \Sigma [(O - E)^2 / E]$

Calculate χ^2:

$$\chi^2 = [(200 - 200)^2 / 200] + [(80 - 50)^2 / 50] + [(10 - 20)^2 / 20] + [(210 - 230)^2 / 230]$$

Step 5: Determine the Degrees of Freedom

Degrees of freedom (df) in this test equal the number of categories minus 1. In this case, df = 4 - 1 = 3.

Step 6: Find the Critical Value or P-value

Using the Chi-squared distribution table or a calculator, you can find the critical value or p-value associated with χ^2 and df.?????

Step 7: Make a Decision

If the calculated χ^2 is greater than the critical value or if the p-value is less than the chosen significance level (α), then you reject the null hypothesis.
If χ^2 is less than the critical value or if the p-value is greater than α, you fail to reject the null hypothesis.

2) Chi-squared Test for Independence (Chi-squared test of association): This test is used to examine the relationship between two categorical variables. In medicine, it's often employed to investigate whether there is a significant association between two factors, such as smoking status and the development of lung cancer. The test is performed on a contingency table to determine if there is a statistically significant association or independence between the two variables.

Null Hypothesis (H0): There is no association between the variables.
Alternative Hypothesis (Ha): There is an association between the variables.

Example: Drug Efficacy and Patient Outcomes
Suppose a pharmaceutical company has developed a new drug designed to treat a specific medical condition, and they want to determine if the drug's efficacy is associated with the age group of patients. They have conducted a clinical trial involving 200 patients, and they collect data on whether each patient's condition improved or not after taking the drug. Additionally, they categorize patients into two age groups: "Young" (under 40 years old) and "Old" (40 years old and above).

The data collected is organized in a contingency table:

Patient Outcome

Age Group	Improved	Not Improved
Young (Under 40)	60	40
Old (40 and above)	45	55

In this contingency table:

The pharmaceutical company is interested in determining whether there is a statistically significant association between the age group of patients and their treatment outcomes (improved or not improved) when using this new drug.???

To analyze this association, they can perform a Chi-squared Test for Independence.

3) Chi-squared Test for Homogeneity: This test is used when you have two or more independent samples, and you want to determine whether the proportions of certain categorical outcomes are similar across these samples. For example, it can be used to compare the success rates of different treatments for a particular medical condition.

Null Hypothesis (H0): The proportions are the same across all groups.
Alternative Hypothesis (Ha): The proportions are different across at least one group.

Example: Comparing Treatment Outcomes in Different Hospitals

Imagine a study that aims to compare the success rates of a specific surgical procedure for a particular medical condition across three different hospitals. The researchers want to determine if there is a significant difference in the success rates among these hospitals. They collect data on the outcomes of the surgical procedure for 300 patients (100 patients from each hospital) and categorize the outcomes into "Successful" and "Unsuccessful.“??????

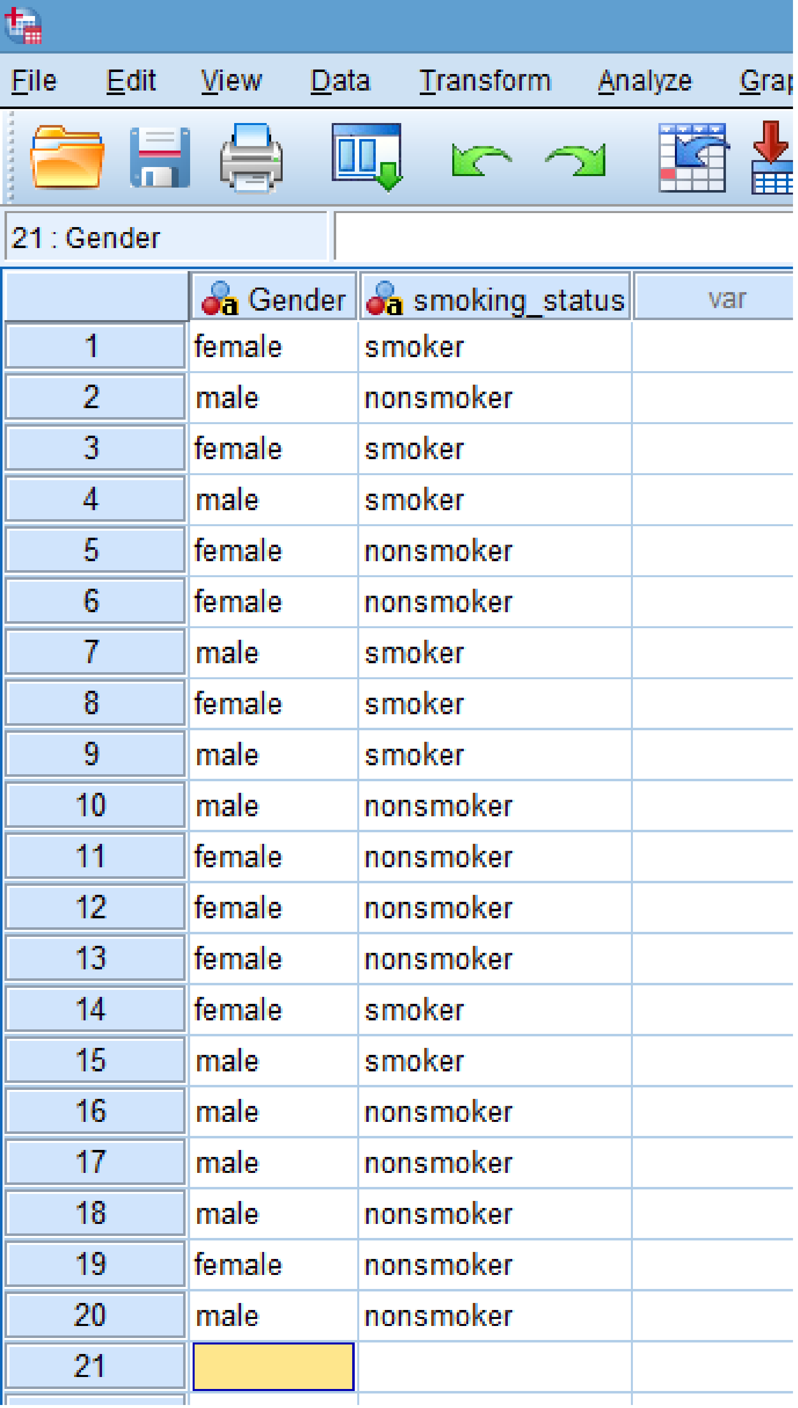
Here's the data organized in a contingency table:

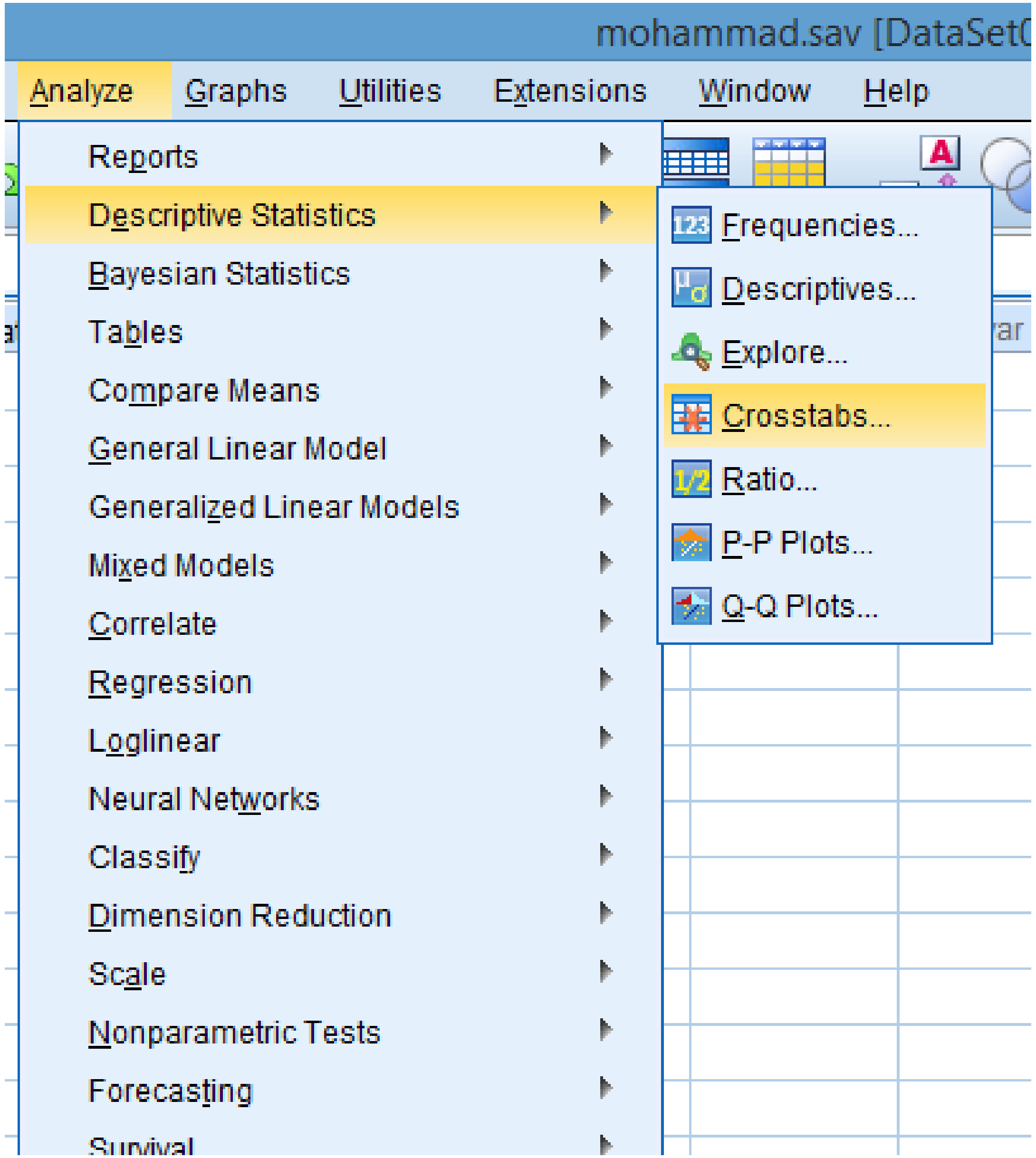
Surgical Outcome

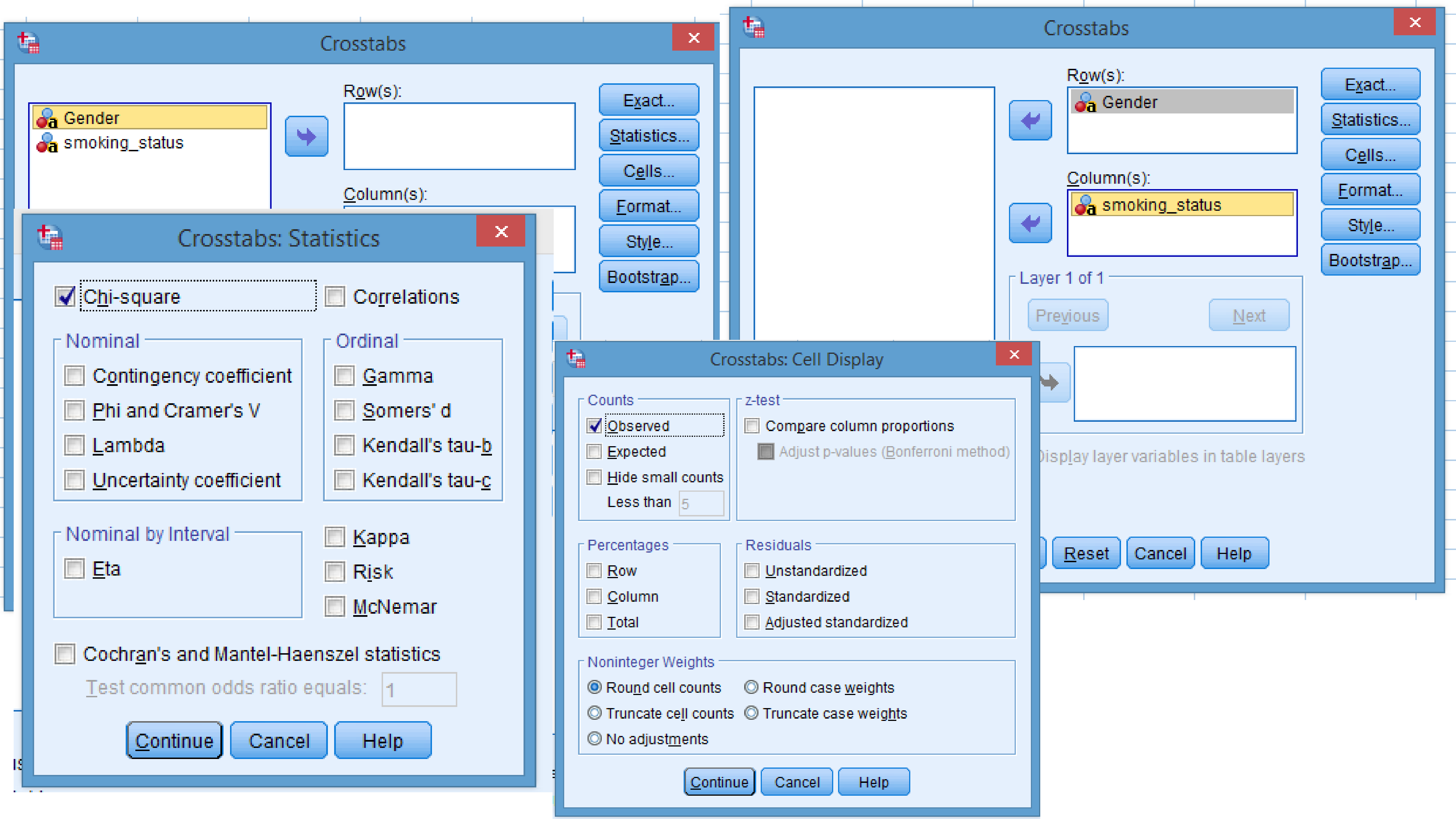
Hospital	Successful	Unsuccessful
Hospital A	75	25
Hospital B	85	15
Hospital C	60	40

In this contingency table:

The researchers want to determine if the proportion of successful outcomes is the same across all three hospitals or if there is a significant difference. To do this, they can use the Chi-squared Test for Homogeneity.


File Edit View Data Transform Analyze Grap
21 : Gender
Gender smoking_status var
1 female smoker
2 male nonsmoker
3 female smoker
4 male smoker
5 female nonsmoker
6 female nonsmoker
7 male smoker
8 female smoker
9 male smoker
10 male nonsmoker
11 female nonsmoker
12 female nonsmoker
13 female nonsmoker
14 female smoker
15 male smoker
16 male nonsmoker
17 male nonsmoker
18 male nonsmoker
19 female nonsmoker
20 male nonsmoker
21


mohammad.sav [DataSet
Analyze Graphs Utilities Extensions Window Help
Reports
Descriptive Statistics
Bayesian Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Frequencies...
Descriptives...
Explore...
Crosstabs...
Ratio...
P-P Plots...
Q-Q Plots...


Crosstabs
Gender
smoking_status
Row(s):
Column(s):
Exact...
Statistics...
Cells...
Format...
Style...
Bootstrap...
Crosstabs: Statistics
Chi-square
Correlations
Nominal
Contingency coefficient
Phi and Cramer's V
Lambda
Uncertainty coefficient
Ordinal
Gamma
Somers' d
Kendall's tau-b
Kendall's tau-c
Nominal by Interval
Eta
Kappa
Risk
McNemar
Cochran's and Mantel-Haenszel statistics
Test common odds ratio equals:
1
Continue
Cancel
Help
Crosstabs
Row(s):
Gender
Column(s):
smoking_status
Layer 1 of 1
Previous
Next
Display layer variables in table layers
Reset
Cancel
Help
Exact...
Statistics...
Cells...
Format...
Style...
Bootstrap...
Crosstabs: Cell Display
Counts
Observed
Expected
Hide small counts
Less than
5
z-test
Compare column proportions
Adjust p-values (Bonferroni method)
Percentages
Row
Column
Total
Residuals
Unstandardized
Standardized
Adjusted standardized
Noninteger Weights
Round cell counts
Round case weights
Truncate cell counts
Truncate case weights
No adjustments
Continue
Cancel
Help

Output9 [Document9] - IBM SPSS Statistics Viewer*

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Output
- Log
- Crosstabs
  - Title
  - Notes
  - Case Processing
  - gender * smoking
  - Chi-Square Tests

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
gender * smoking	20	100.0%	0	0.0%	20	100.0%

gender * smoking Crosstabulation

Count

		smoking		
		nonsmoker	smoker	Total
gender	female	6	4	10
	male	6	4	10
Total		12	8	20

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.000[a]	1	1.000		
Continuity Correction[b]	.000	1	1.000		
Likelihood Ratio	.000	1	1.000		
Fisher's Exact Test				1.000	.675
N of Valid Cases	20				

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 4.00.

b. Computed only for a 2x2 table

IBM SPSS Statistics Processor is ready | Unicode:ON

03:30 م
٢٠٢٣/١٠/٢٨