MEDICAL UNIVERSITY – PLEVEN
FACULTY OF PUBLIC HEALTH

DEPARTMENT OF PUBLIC HEALTH SCIENCES
CENTRE FOR DISTANT LEARNING

LECTURE No7

HYPOTHESIS TESTING. PARAMETRIC AND NON-PARAMETRIC TESTS

Assoc. Prof. Gena Grancharova, MD, PhD

Plan of the lecture

1. Introduction and logic of hypothesis testing
2. Basic concepts
3. Basic steps in hypothesis testing
4. Parametric and non-parametric tests

1. INTRODUCTION AND LOGIC OF HYPOTHESIS TESTING

What is statistical hypothesis testing?

It is simply the process of decision making.

Suppose that a physician researcher hypothesized that cancer patients' participation in a stress management programme would result in lower anxiety.

Two groups were observed. The first group consisted of 25 patients as a control group (they didn't participate in a stress management programme and 25 subjects in the experimental group that were subjected to a stress management programme.

The researcher found that the mean anxiety level for the experimental group was 15.8 and that of the control group is 17.5.

Should the researcher conclude that the hypothesis stated has been supported?

In fact, some group differences are observed and they were in a predicted direction, but the results might simply be the result of sampling fluctuations.

Statistical hypothesis testing allows researchers to make objective decisions concerning the results of their studies. Scientists need such a mechanism for helping them to decide which outcomes are likely to reflect only chance differences between sample groups and which are likely to reflect true population differences.

To answer such question a researcher should use ***tests of significance.***

Tests of significance are standard statistical procedures for drawing inferences from sample estimates about unknown population parameters. Sample estimates are never exact, being subject to sampling errors. Thus, in the design of any medical research, attempts are made to reduce these sampling errors.

Tests of significance allow the researchers to decide whether the sample estimates, or the differences between estimates, are within their normal biological variation, commonly called **variability due to chance or chance variation.**

So, any time when a difference is observed it is important to answer the question **whether such a difference has occurred by chance alone or it is due to some other causes.**

The possible causes of observed differences may be related:

— to chance variation;

— to the factor under study;

— to the other "real" factors;

— to some "spurious" factors, such as bias and non-comparability.

THE LOGIC OF HYPOTHESIS TESTING

Hypothesis testing is the process of deciding statistically whether the findings of an investigation reflect chance or 'real' effects at a given level of probability.

The mathematical procedures for hypothesis testing are based on the application of probability theory. Because of this, decision errors in hypothesis testing cannot be entirely eliminated.

However, the process allows to specify the probability level at which we can claim that the data obtained in an investigation support experimental hypotheses.

2. BASIC CONCEPTS

= Null hypothesis

= Alternative hypothesis

- Directional (one-tailed)
- Non-directional (two-tailed)

= Statistical significance

= Statistical test

- Parametric tests
- Non-parametric tests

= Type I and type II errors

Null hypothesis - H_0

The procedures used in testing hypothesis are based on negative inference. This logic seems somewhat unusual to students and to beginning researchers.

At the previous example, the researcher had tested the effectiveness of a programme to reduce stress and anxiety in cancer patients and he found a difference in experimental and control groups.

There two possible explanations for this outcome:

1. the experimental treatment was successful in reducing patients' anxiety;

2. the differences may result to chance factors (such as differences in anxiety levels of the two groups before the treatment.

The first explanation corresponds to the researcher's scientific hypothesis.

The second explanation corresponds to the null hypothesis.

Null hypothesis - H_0 is a statement that there is no actual relationship between dependent and independent variables *(level of anxiety and stress management programme) and that any observed relationship is only a function of chance or sampling fluctuations.*

Alternative hypothesis – H_1 or H_A (experimental hypothesis) – it's the hypothesis for which the researcher is trying to gain support through statistical analysis, by rejecting the null hypothesis.

H_1 states that there is a difference between the groups or a relationship between dependent and independent variables.

There are two types of experimental hypothesis:

- **Directional hypothesis or one-tailed hypothesis**

- **Non-directional hypothesis or two- tailed hypothesis**

Directional hypothesis (one-tailed) – it asserts that differences between groups in the data will occur in a particular direction, *e.g. smokes die younger than non-smokers*.

Non-directional hypothesis (two-tailed) – it asserts that there are differences between groups in the data but with no direction specified, *e.g. smokers and non-smokers have different life expectancies.*

The statistician normally poses the null hypothesis and then tests it statistically.

If it is rejected, then the alternative hypothesis (there is a difference between two groups or a relationship between variables) is accepted.

Statistical significance (P) – it's the probability over which H_0 is accepted to be true and below which H_0 is rejected.

P for H_0 + P for H_1 = 1 = 100%

When P > 0.05 - H_0 is true.
The conclusion is that **there is no difference between two groups** or a relationship between variables (if some difference is observed it is due to chance).

When P < 0.05 - H_0 is false, it is rejected and H_1 is accepted. The conclusion is that **there is a real difference between the groups** or a relationship between variables.

Statistical test – it's a statistic calculated from the sample data and its value is used to decide whether H_0 is to be accepted or rejected.

Two types of statistical tests:

Parametric tests – suitable for the analysis of interval or ratio data.

Non-parametric tests – suitable for the analysis of nominal or ordinal data.

Any decision made on probabilistic basis might be erroneous.

Two types of elementary decision errors can be identified - **Type I and Type II errors.**

Type I error (α) involves mistakenly rejecting H_0, when it is true.

Type II error (β) involves mistakenly accepting H_o when it is false.

Real situation	Decision	
	H_0 is rejected	H_0 is accepted
H_0 is true	Type 1 error	Right decision
H_1 is true	Right decision	Type II error

We can minimize the Type I error by setting an acceptable level for α.

In scientific research, editors of most scientific journals require that α should be set at 0.05 or less.

How can we minimize Type II error?

1. By increasing the sample size, n.

2. By reducing the variability of measurements (s), either by increasing accuracy or by using samples which are not highly variable for the measurement producing the data.

3. By using a directional H_1, on the basis of previous evidence about the nature of the effect.

4. By setting a less demanding α, type I error rate.

There is a relationship between α and β, such that the smaller α, the greater β.

3. BASIC STEPS IN HYPOTHESIS TESTING

1. State the null hypothesis (H_0), which claims that any differences in the data were just due to chance: the independent variable has no effect on the dependent variable, or that any difference among groups is due to random effects.

2. State the alternative hypothesis (H_1) - the prediction which we intend to evaluate.

H_1 claims that the results are 'real' or 'significant': the independent variable influenced the dependent variable, or that there is a real difference among groups.

3. Decide the type of H_1 – directional (one-tailed) or non-directional (two-tailed).

4. State the level of significance (the decision level)

There are two mutually exclusive hypotheses (H_1 and H_0) competing to explain the results of an investigation.

Hypothesis testing, or statistical decision making, involves establishing the probability of H_0 being true.

If this probability is very small, we are in a position to reject the H_0.

The conventional levels of significance:

- "significant" for $p < 0.05$;

- "highly significant" for $p < 0.01$;

- "not significant" for $p > 0.05$ or $p = 0.05$.

If the probability of H_0 being true is less than 0.05 or 0.01, we can reject H_0.

5. Choose the test statistic

Statistical test – a statistic calculated from the sample data whose value is used to decide whether H_0 is to be accepted or rejected.

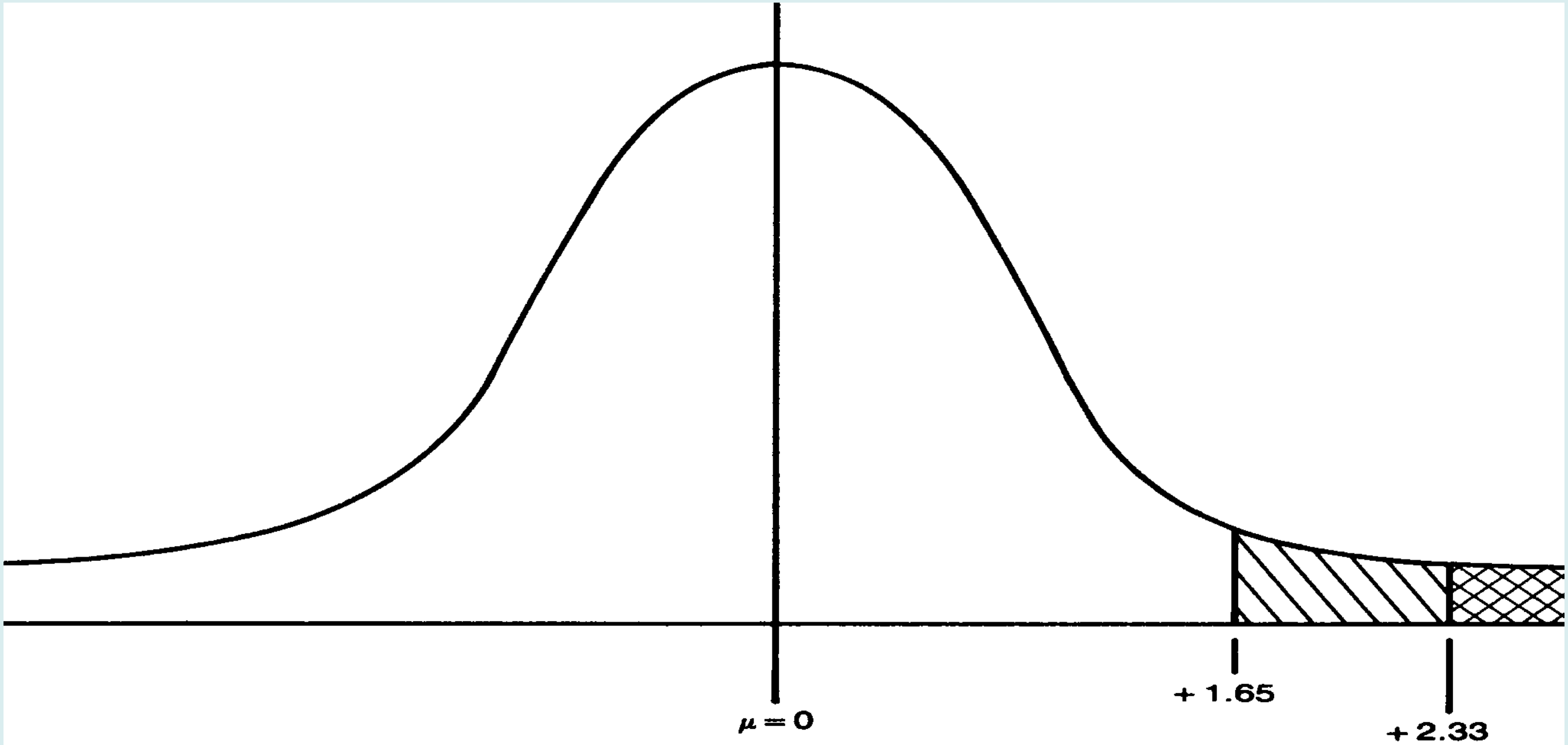
Parametric tests – for the analysis of interval or ratio data, e.g. t-test

Non-parametric tests – for the analysis of nominal or ordinal data, e.g. χ^2

6. Compute the numerical value of the test statistic from the observed data to decide the probability of H_0 being true. That is, we assume H_0 is true, and calculate the probability of the outcome of the investigation being due to chance alone.

7. Compare the calculated value of the test statistic with tabulated critical values in appropriate standard distribution tables at a specified probability level of significance.

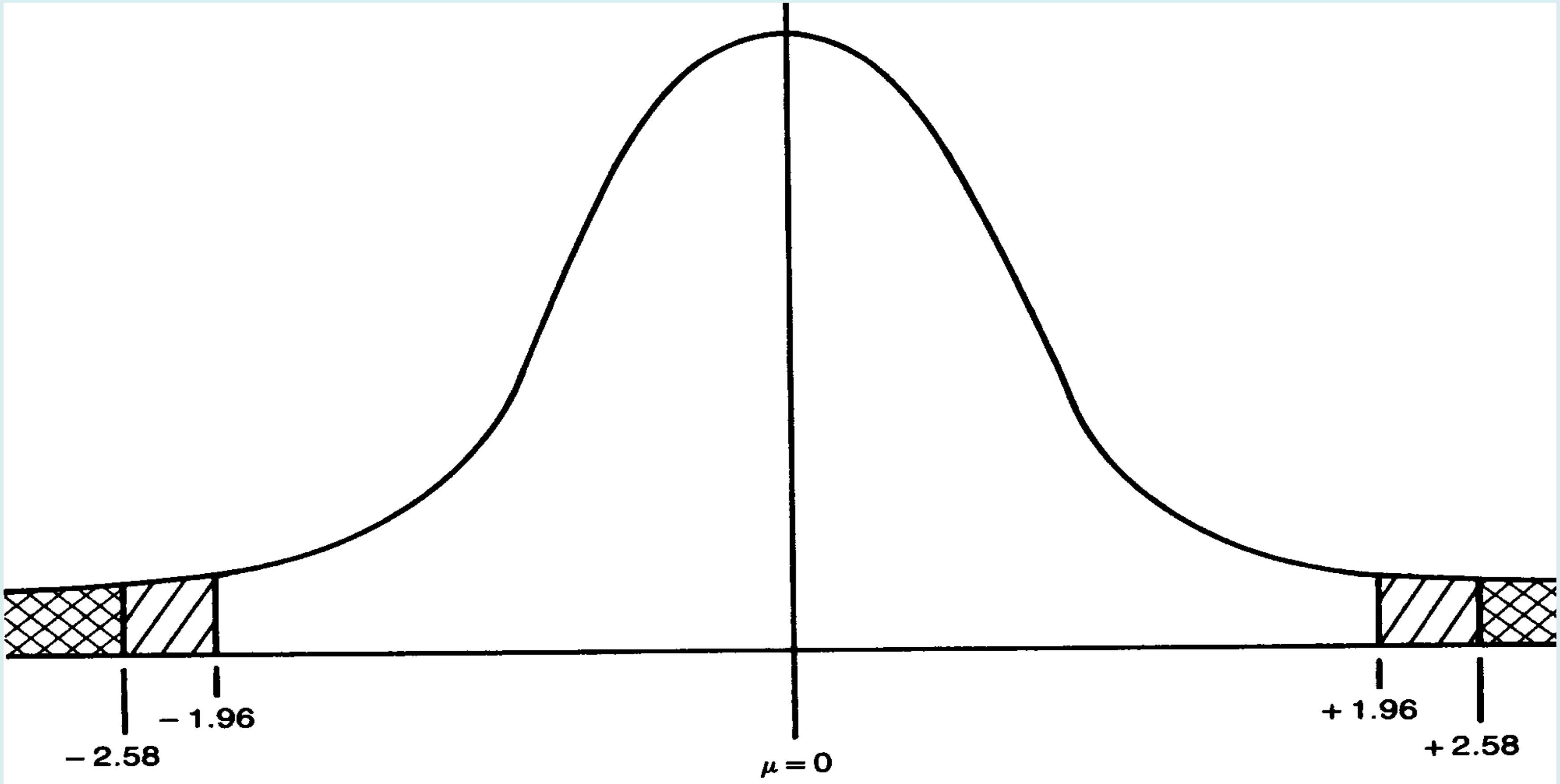
There may be two types of tests:
One-tailed test - a statistical test where a difference between two groups is tested in a particular direction of the difference, e.g. to test a **directional hypothesis** – when not only the significance of differences is tested but also the direction of these differences is determined. In other words, the critical area for one-sided test is a series of values that are less or higher that the critical value of the test.



One-sided test of significance in a normal curve

Two-tailed test – a statistical test where a difference between two groups is tested without reference to the expected direction of the difference, e.g. **for non-directional hypothesis.**

The critical area for two-sided test is a series of values that are less that the first critical value of the test and a series of values that are higher than the second critical value of test.



Two-sided test of significance in a normal curve

Level of significance for H0 in two-tailed test							
	P=0.1	0.05	0.02	0.01	0.005	0.002	0.001
Level of significance for H0 in one-tailed test							
K (df)↓	P=0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
3	2.353	3.182	4.541	5.841	7.453	10.214	12.924
4	2.132	2.776	3.747	4.604	5.498	7.173	8.610
5	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	1.943	2-447	3.143	3.707	4.317	5.208	5.959
7	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	1.812	2.228	2.764	3.169	3.581	4.144	4.587
15	1.753	2.131	2.602	2.947	3.286	3.733	4.073
20	1.725	2.086	2.528	2.845	3.153	3.552	3.850
25	1.708	2.060	2.485	2.787	3.078	3.450	3.725
30	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	1.671	2.000	2.390	2.660	2.915	3.232	3.460
120	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	1.645	1.960	2.326	2.576	2.807	3.090	3.291

8. Decide whether or not to reject H_0 according to the p-value.

If $p \geq 0{,}05$ – H_0 is true (it is accepted).

If $p < 0.05$ – H_0 is rejected and H_1 is accepted.

Decision rules

1. If the magnitude of the obtained value of the statistic exceeds the critical value, H_o is rejected.

obtained test values ≥ critical values – reject H_0

In other words, When $p<0.05$ in two-tailed test or $p<0.025$ in one tailed test - H_0 is rejected and H_1 is accepted, e.g. there is a significant difference.
In this case, the investigator concludes that the data supported the differences predicted by the alternative hypothesis (at the level of significance).

Decision rules

2. If the obtained value of the statistic calculated is less than the critical value, H_0 is accepted.

obtained test values < critical values – retain H_0

In other words, if p-value for H_0 is $p>0.05$ in two-tailed test ($p>0.025$ in one-tailed test) then H_0 is true, e.g. there is no significant difference; the difference observed is due to chance.

Interpretation of p-values

- Statistical significance versus medical importance or significance

- a statistically significant difference but of no clinical importance;

- a non-statistically significant observation but with the results pointing to a possible clinical or medical importance.

- Role of sample size in determining statistical significance

4. PARAMETRIC AND NON-PARAMETRIC METHODS FOR HYPOTHESIS TESTING

Both parametric and non-parametric methods for hypothesis testing are based on the **same logic** and use the **same methodological steps of work.**

The difference is only in the last **steps 6th and 7th** concerning **the approaches in calculation of appropriate tests** and using **different tables of critical values** to determine the level of probability (the level of statistical significance).

4. 1. PARAMETRIC METHODS FOR HYPOTHESIS TESTING

Parametric tests are used when the data are measured on interval or ratio scale and a normal distribution is assumed.

The most widely used **tests** are:

= t-test (paired or unpaired);

= **ANOVA** (analysis of variances) - one-way non-repeated, repeated, two-way, three-way);

= **linear regression.**

PAIRED T-TEST (FOR INDEPENDENT AND DEPENDENT SAMPLES)

The paired t-test is the most commonly used method to evaluate the differences in means between two groups.

The groups compared can be:
- **independent** (e.g., the means of blood pressure of patients who were given a drug vs. a control group who received a placebo);

- **dependent** (e.g., the means of blood pressure of patients "before" vs. "after" they received a drug).

How the paired t-test can be applied?

The same steps in hypothesis testing described in the first part of the lecture should be followed for paired t-test.

1. State the null hypothesis (H_0), which claims that any differences in the data were just due to chance: the independent variable has no effect on the dependent variable, or that any difference among groups is due to random effects.

2. State the alternative hypothesis (H_1) - the prediction which we intend to evaluate.

H_1 claims that the results are 'real' or 'significant': the independent variable influenced the dependent variable, or that there is a real difference among groups.

3. Decide the type of H_1:

= directional (one-tailed) or

= non-directional (two-tailed).

Examples of directional and non-directional hypotheses

For **directional hypothesis** we use **one-sided or one-tailed test.**

For **non-directional hypothesis** we use **two-sided or two-tailed test.**

4. State the level of significance (the decision level)

There are two mutually exclusive hypotheses (H_1 and H_0) competing to explain the results of an investigation.

Hypothesis testing, or statistical decision making, involves establishing the probability of H_0 being true.

If this probability is very small, we are in a position to reject the H_0.

The conventional levels of significance:

- "significant" for $p < 0.05$;

- "highly significant" for $p < 0.01$;

- "not significant" for $p > 0.05$ or $p = 0.05$.

This means if the probability of H_0, being true is less than 0.05 or 0.01, we can reject H_0 and accept the alternative hypothesis H_1.

5. Choose the test statistic

Statistical test – a statistic calculated from the sample data whose value is used to decide whether H_0 is to be accepted or rejected.

Parametric tests – for the analysis of interval or ratio data, e.g. t-test

Non-parametric tests – for the analysis of nominal or ordinal data, e.g. chi-square.

6. Compute the numerical value of the test statistic from the observed data to decide the probability of H_0 being true. That is, we assume H_0 is true, and calculate the probability of the outcome of the investigation being due to chance alone.

Calculation of t-tests

Besides using statistical software packages, such as SPSS, we can simply calculate t-criterion.

For independent samples with different variances t can be calculated by the following formula:

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where: $|\bar{x}_1 - \bar{x}_2|$ is the absolute difference of the means in both groups

s_1 and **s_2** – standard deviations in both groups

n_1 и n_2 – number of cases in both groups

If the calculated statistics are **proportions or rates** then the t-test is calculated as:

$$t = \frac{|p_1 - p_2|}{\sqrt{\dfrac{p_1 \times q_1}{n_1} + \dfrac{p_2 \times q_2}{n_2}}}$$

where:

p_1 and p_2 – proportions in both groups

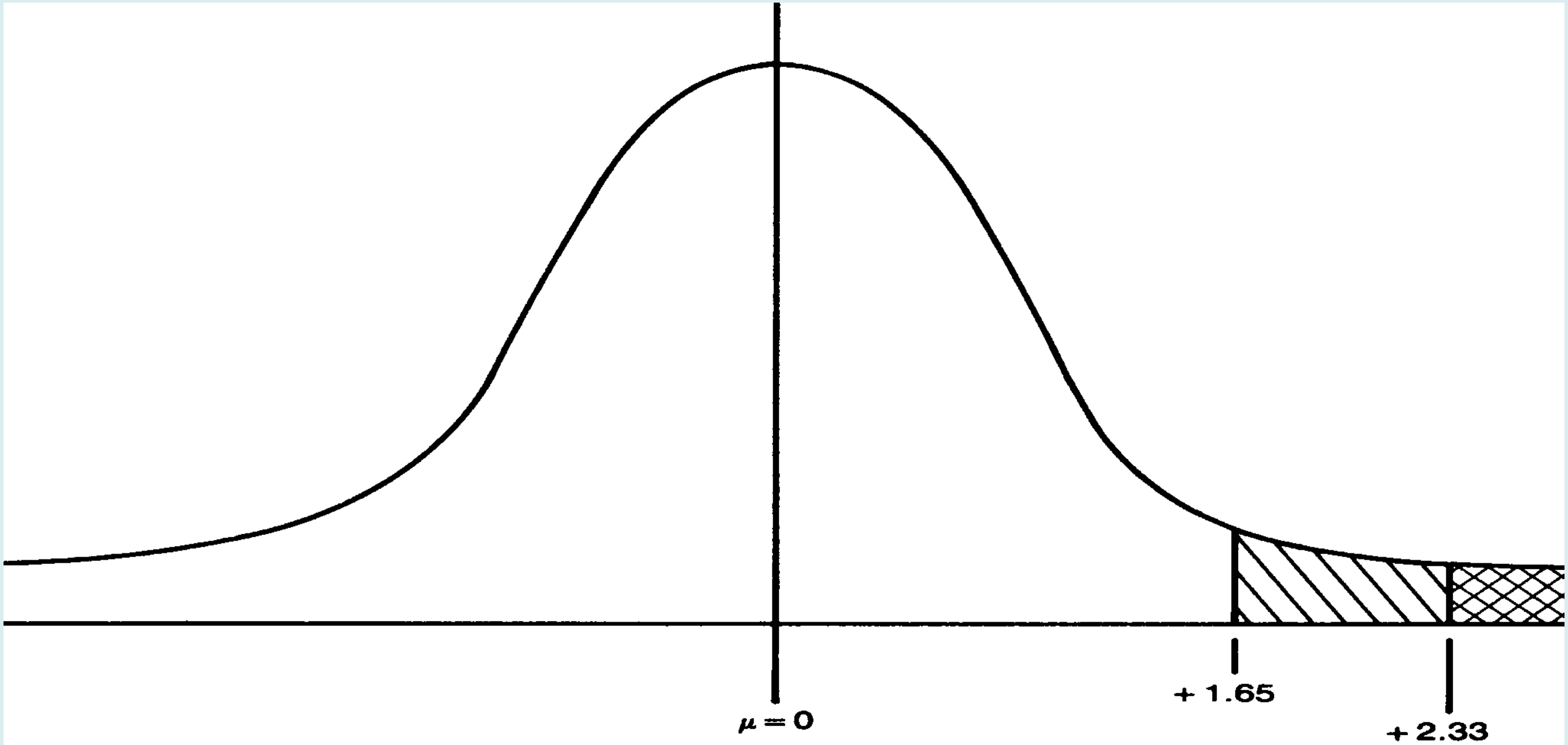
q_1 and q_2 – values to be added to proportions in both groups to come to 1, 100, 1000, etc.

n_1 and n_2 – number of cases in both groups

7. Compare the calculated value of the test statistic with tabulated critical values in appropriate standard distribution tables at a specified probability level of significance.

Tables of critical values of t-test provide opportunity for two types of tests:

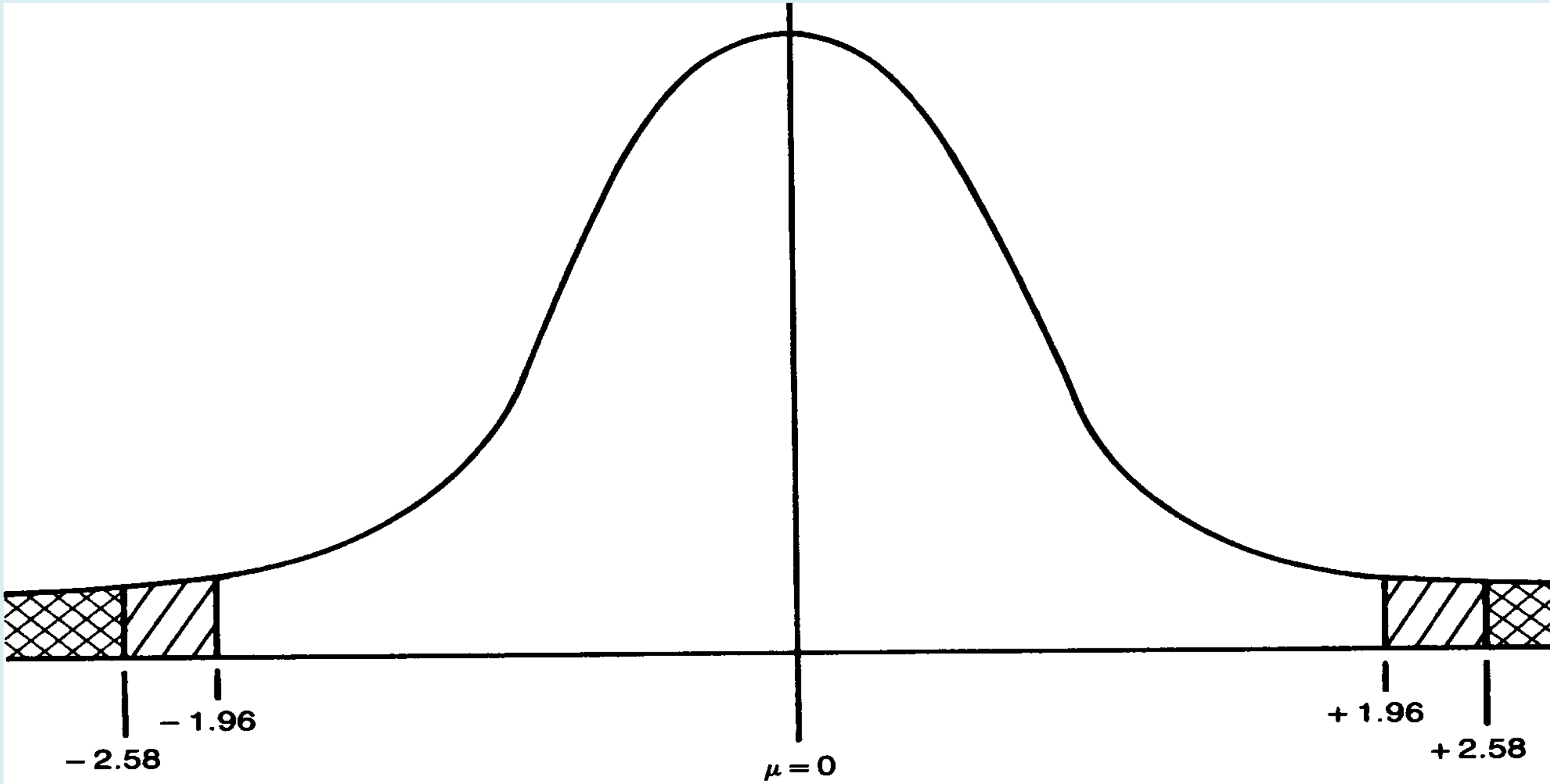
One-tailed test - a statistical test where a difference between two groups is tested in a particular direction of the difference, e.g. to test a **directional hypothesis.**



One-sided test of significance in a normal curve

Two-tailed test – a statistical test where a difference between two groups is tested without reference to the expected direction of the difference, e.g. ***for non-directional hypothesis.***

The critical area for two-sided test is a series of values that are less that the first critical value of the test and a series of values that are higher than the second critical value of test.



Two-sided test of significance in a normal curve

Level of significance for H0 in two-tailed test							
	P=0.1	**0.05**	**0.02**	**0.01**	**0.005**	**0.002**	**0.001**
Level of significance for H0 in one-tailed test							
K (df)↓	**P=0.05**	**0.025**	**0.01**	**0.005**	**0.0025**	**0.001**	**0.0005**
3	2.353	3.182	4.541	5.841	7.453	10.214	12.924
4	2.132	2.776	3.747	4.604	5.498	7.173	8.610
5	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	1.943	2-447	3.143	3.707	4.317	5.208	5.959
7	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	1.812	2.228	2.764	3.169	3.581	4.144	4.587
15	1.753	2.131	2.602	2.947	3.286	3.733	4.073
20	1.725	2.086	2.528	2.845	3.153	3.552	3.850
25	1.708	2.060	2.485	2.787	3.078	3.450	3.725
30	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	1.671	2.000	2.390	2.660	2.915	3.232	3.460
120	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	1.645	1.960	2.326	2.576	2.807	3.090	3.291

8. Decide whether or not to reject H_0 according to the p-value.

If $p>0,05$ – H_0 is true (it is accepted).

If $p<0.05$ – H_0 is rejected and H_1 is accepted.

4.2. NON-PARAMETRIC TESTS

Non-parametric tests are used with nominal or ordinal variables. They do not require a distribution to meet the required assumptions to be analyzed (especially if the data is not normally distributed). Due to such a reason, they are sometimes referred to as **distribution-free tests,** **e.g.** **they can be applied to any type of distributions.** **That's why they are very commonly used.**

Non-parametric tests do not substitute the parametric tests but they serve as their alternative. Thus, parametric tests often have nonparametric equivalents

Nonparametric Tests	Parametric Tests
Mann-Whitney U Test	Independent Samples T-test
Wilcoxon Signed Rank Test	Paired Samples T-test
Kruskal-Wallis Test	One-way ANOVA
Chi-squared Test	

The most widely used non-parametric test is the **Chi-Square test of independence.**

The χ^2 (chi-square) test

χ^2 is appropriate for statistical analysis when:

1. Variables are categorical - measured on a nominal or ordinal scale.
2. Measurements were of independent subjects.

Chi-Square test of independence is used to determine if there is a significant relationship between two nominal (categorical) variables. The frequency of each category for one nominal variable is compared across the categories of the second nominal variable.

So, **frequency tables are required to present the observed data.**

Frequency tables of two variables presented simultaneously are called **contingency tables,** constructed by listing all the levels of one variable as rows **in a table** and the levels of the other variables as columns, then finding the joint or cell **frequency** for each cell.

Two types of contingency tables:
= 2x2 (each variable has 2 categories**);**
= multiple contingency table (at least one of the variables has more than two categories).

The chi-square test compares the observed vs. the expected frequencies.

Observed Frequencies are counts made from experimental data, e.g. actually observed and measured data.

Expected frequencies are counts related to the probability of the null hypothesis to be true. For the chi-squared test to give meaningful results the expected frequency for each cell in the 2x2 contingency table is required to be **at least 5.**

The chi-square is given by the formula:

$$\chi^2 = \Sigma \frac{(f_o - f_e)^2}{f_e}, \text{ where:}$$

f_o – observed frequency for a given category

f_e - expected frequency for a given category if H_o is true

Example: In 10-year longitudinal cohort study the frequency of chronic obstructive pulmonary disease (COPD) was studied among two groups: smokers and non-smokers. The smoking here is an independent variable (the factor whose impact is studied), and the occurrence of COPD is an outcome (dependent variable).

	With COPD	Without COPD	TOTAL
Smokers	100 (75) a	400 (425) b	500 – a + b
Non-smokers	50 (75) c	450 (425) d	500 – c + d
TOTAL	150 a + c	850 b + d	1000

Steps in hypothesis testing:

1. H_o - there is no difference

2. H_1 – there is a difference in COPD in smokers and non-smokers

3. Defining the expected frequencies.

Expected frequencies are calculated by the formula:

$$\frac{\text{total row x total column}}{\text{grand total}}$$

The first theoretical value is equal to (150.500):1000=75. We put this value in brackets in the same table cell. The other three theoretical values add the results in the summary row and column.

4. Determine the degree of freedom –
df = (2-1).(2-1) = 1 in 2x2 tables
df = (r-1). (c-1) in multiple tables, where r is a number of rows and c – number of columns

5. Calculation of χ^2

$$\chi^2 = \frac{(100-75)^2}{75} + \frac{(400-425)^2}{425} + \frac{(50-75)^2}{75} + \frac{(450-425)^2}{425} = 19.6$$

We can come to the same result using the the other formula:

$$\chi^2 = \frac{N\,(ad - bc)^2}{(a+b).(c+d).(a+c).(b+d)} = \frac{1000.(100.450 - 50.4000)2}{500.500.150.850} = 19.6$$

6. Comparing χ^2 with the table of critical values

7. Conclusion

CRITICAL VALUES OF χ^2

$H_0 \rightarrow$	P=0.100	0.050	0.025	0.010	0.005	0.001
$H_1 \rightarrow$	1-P=0.900	0.950	0.975	0.990	0.995	0.999
K (df) ↓ 1	2.71	3.84	5.02	6.63	7.88	10.83
2	4.61	5.99	7.38	9.21	10.60	13.82
3	6.25	7.81	9.35	11.34	12.84	16.27
4	7.78	9.49	11.14	13.28	14.86	18.47
5	9.24	11.07	12.83	15.09	16.75	20.52
6	10.64	12.59	14.45	16.81	18.55	22.46
7	12.02	14.07	16.01	18.48	20.28	24.32
8	13.36	15.51	17.53	20.09	21.96	26.13
9	14.68	16.92	19.02	21.67	23.59	27.88
10	15.99	18.31	20.48	23.21	25.19	29.59
15	22.31	25.00	27.49	30.58	32.80	37.70
20	28.41	31.41	34.17	37.57	40.00	45.32
25	34.38	37.65	40.65	44.31	46.93	52.62
30	40.26	43.77	46.98	50.89	53.67	59.70
40	51.81	55.76	59.34	63.69	66.77	73.40
50	63.17	67.50	71.42	76.15	79.49	86.66
60	74.40	79.08	83.30	88.38	91.95	99.61

Comparing the obtained value of chi-square 19.6 at the degree of freedom 1, we find that it is much higher than the critical value corresponding to $p=0.001$. This mean that in this case $p<0.001$. The null hypothesis should be rejected and the conclusion is **that there is a highly significant difference in occurrence of COPD in smokers and non-smokers.**

TEST EXAMPLES

1. The null hypothesis states that there is no difference in the results of the groups compared.

A. True

B. False

2. The alternative hypothesis states that there is an effect or difference in the results.

A. True

B. False

3. The level of significance can be viewed as the level of confidence with witch the final decision is supported.

A. True

B. False

4. One-tailed test will determine
A. If the two extreme values (min or max) of the sample need to be rejected
B. If the hypothesis has one or possible two conclusions
C. If the region of rejection is located in one tail of the distribution

5. Two-tailed test will determine
A. If the two extreme values (min or max) of the sample need to be rejected
B. If the hypothesis has one or possible two conclusions
C. If the region of rejection is located in two tails of the distribution

6. In hypothesis testing Type II error is committed when:
A. We reject the null hypothesis whilst the alternative hypothesis is true
B. We reject a null hypothesis when it is true
C. We accept a null hypothesis when it is not true

7. In hypothesis testing Type I error is committed when:
A. We reject the null hypothesis whilst the alternative hypothesis is true
B. We reject a null hypothesis when it is true
C. We accept a null hypothesis when it is not true

8. Contingency tables and degrees of freedom are the key elements of performing the chi-square test.
A. True
B. False

9. For the chi-square test to be effective, the expected value for each cell in the contingency table has to be at least:
A. 3
B. 5
C. 10

10. By taking a level of significance of 0.05 for the null hypothesis it is the same as saying:

A. We are 5% confident the results have not occurred by chance
B. We are 95% confident that the results have not occurred by chance
C. We are 95% confident that the results have occurred by chance

11. Two types of errors associated with hypothesis testing are Type I and Type II. Type II error is committed when:

A. We reject the null hypothesis whilst the alternative hypothesis is true
B. We reject a null hypothesis when it is true
C. We accept a null hypothesis when the alternative hypothesis is true

12. Two types of errors associated with hypothesis testing are Type I and Type II. Type I error is committed when:

A. We reject the null hypothesis whilst the alternative hypothesis is true
B. We reject a null hypothesis when it is true
C. We accept a null hypothesis when the alternative hypothesis is true

13. Case study: We are interested in investigating whether a novel drug is effective as a weight reducing agent. 30 clinically obese men with a body mass index (BMI) between 30 and 35 are randomly allocated to receive either the drug or a placebo. Each takes the relevant preparation once a day for two months, whilst eating a normal diet, and his BMI is measured at the end of the period.

Select the statement which would provide an appropriate null hypothesis for the study.

A. At the end of the two-month period, the mean BMI in the placebo group is greater than that of the drug group.

B. At the end of the two-month period, the mean BMI in the placebo group is less than that of the drug group.

C. The mean change in BMI from baseline to two months is equal in the two groups.

14. Case study:
Subjects from families with genetic disorders were asked whether they had encountered problems when applying for life insurance. A sample from the general population was also asked the same question. About 33% of respondents in the study group reported problems compared with only 5% of the general population. This difference was significant at the 0.01% level. Select the statement which you believe to be true.

A. The differences between the two groups of families are likely to have occurred by chance.

B. A suitable null hypothesis would be that subjects from families with genetic disorders are more likely to experience problems when applying for life insurance than those from the general population.

C. We can reject the null hypothesis and accept the alternative hypothesis.

15. The Paired Samples *t* Test is appropriate when the variable of interest is binary (binominal, dichotomous).

A. True B. False

16. The Paired Samples *t* Test is appropriate when the variable of interest is numerical (measured on interval or ratio scale).

A. True B. False

17. The Paired Samples *t* Test is appropriate for comparing two means that represent measurement taken under two different conditions (like in independent samples - control and experimental groups).

A. True B. False

18. The Paired Samples *t* Test is appropriate for comparing two means that represent measurement taken in one sample at two different times (like pre-test and post-test with an intervention administered between the two time points).

A. True B. False

19. The critical value of a statistic is the value which cuts off the region for rejecting of the null hypothesis H_o.

A. True B. False

20. If the critical value of a statistic (t or chi-square) is less than the obtained or calculated value, then we can reject the null hypothesis H_o and accept the alternative hypothesis H_1.

A. True B. False

21. If the critical value of a statistic (t or chi-square) is higher than the obtained or calculated value, then we can accept the null hypothesis H_o.

A. True B. False

22. If H_o is true and we accept it, we have made a right decision.

A. True B. False

23. If we reject H_o, then we are in a position to accept H_1.
A. True B. False

24. If H_o is false and we reject it, we have made a Type II error.
A. True B. False

25. If H_o is false and we fail to reject it, we have made a Type II error.
A. True* B. False

26. Determine the statistical significance between the average weight of new-born males (3400 g) and new-born females (3250) if the degree of freedom is df (k) = ∞ and t = 2.85.
A. There is no a significant difference between the means
B. There is a significant difference between the means

27. An investigator is interested in the variables affecting smoking in a college population. The smoking-on-campus study is undertaken, involving 200 males (100 smokers) and 200 females (50 smokers). The results are presented in 2x2 table. The value of chi-square of independence is 36.6. What conclusion could be made?

A. The difference in smoking between males and females is due to chance

B. The difference is not statistically significant

C. The difference is statistically significant

28. An investigator is interested in the variables affecting smoking in a college population. The smoking-on-campus study is undertaken, involving 200 males (100 smokers) and 200 females (50 smokers). The results are presented in 2x2 table. The value of chi-square of independence is 26.6. What is the p-value?

A. $p > 0.05$

B. $p < 0.025$

C. $p < 0.001$

29. What type of data do you need for a chi-square test?
A. Measured on ratio scale
B. Measured on interval scale
C. Measured on nominal scale

30. What does a significant result in a chi-square test imply?
A. There is a significant difference between the distribution of the variables
B. There is a significant relationship between the compared variables
C. Both statements are true

31. What would a chi-square significance value of P > 0.05 suggest?
A. There is no significant difference between the sample and the population
B. There is no significant difference between categories
C. There is a significant relationship between categorical variables

32. The degrees of freedom for the Chi-Square test statistic when testing for independence in a contingency table with 4 rows and 4 columns would be:

A. 12 B. 5 C. 9

33. In general, the expected frequencies per cell in the conduct of a Chi-Square test are those one would:

A. expect to find in a given cell if either the null hypothesis or the alternative hypothesis was actually true

B. expect to find in a given cell if the alternative hypothesis was actually true

C. expect to find in a given cell if the null hypothesis was actually true

34. With a chi-square = 13.28 and df = 4, the difference between the compared groups is:

A. due to chance

B. statistically significant

C. not statistically significant

35 Determine the statistical significance between the average weight of new-borns in rural and urban areas if the degree of freedom is df = 200 and paired t-test = 1.28.

A. There is a significant difference between the means

B. There is no significant difference between the means

C. None of the above

36. A public opinion poll surveyed a simple random sample of voters. Respondents were classified by gender (males or females) and by age (under 50 years and above 50 years). The value of chi-square was 2.67. What is your conclusion about the significance of the difference observed?

A. The differences between the two groups are likely to occur by chance

B. There is no significant difference between the two groups

C. Both B and C are true

37. Given chi-square = 9.6 and degree of freedom df = 6, the difference between the compared groups is:

A. due to chance

B. statistically significant

C. there is not enough information

38. State the level of significance of H_0 with chi-square = 6.2 and df = 2:

A. p $(H_0) < 0.05$

C. p $(H_0) < 0.01$

B. p $(H_0) > 0.05$

39. What is the type of the hypothesis stating that mortality rates from lung cancer in smokers are different from those in non-smokes?

A. Directional

B. One-tailed

C. Non-directional

40. What is the type of the hypothesis stating that mortality rates from lung cancer in smokers are higher from those in non-smokes?

A. Directional
B. Two-tailed
C. Non-directional

41. What is the type of the hypothesis stating that the life expectancy in females is different from that in males?

A. Directional
B. One-tailed
C. Non-directional

42. What is the type of the hypothesis stating that the life expectancy in females is higher than in males?

A. Directional
B. Two-tailed
C. Non-directional

43. A directional research hypothesis (H_1) should be used when there is theoretical justification for the existence of a directional effect in the data.

A. True B. False

44. Select one of the following statements which you believe to be true. The paired t-test is appropriate when:

A. The differences between the pairs of observations are normally distributed.

B. We wish to test the null hypothesis that the mean of the differences between the pairs of observations in the sample is equal to zero.

C. We wish to test the null hypothesis that the median of the differences between the pairs of observations in the population is equal to zero.

45. The decision level (statistical significance) in hypothesis testing is generally set at 0.05 or 0.01.

A. True B. False

46. The closer the observed frequency for each cell is to the expected frequency, the higher the probability of rejecting the null hypothesis H_0 when using chi-square.

A. True B. False

47. A basic assumption for using t is that the samples were drawn from normally distributed population.

A. True B. False

48. A basic assumption of chi-square is that the scores in each cell are independent.

A. True B. False

49. Parametric tests are used to analyse the significance of interval or ratio data.

A. True B. False

50. The use of non-parametric tests depends on the normal distribution of the underlying population.

A. True B. False

51. If the values of expected and observed frequencies are the same for each cell in the contingency table, chi-square will not be statistically significant.

A. True B. False

52. The degree of freedom in 2x2 contingency tables is always equal to 1.

A. True B. False

53. The degree of freedom in multiple contingency tables is always greater than 1.

A. True B. False

54. The degree of freedom in multiple contingency tables depends on the number of categories in rows and in columns and is calculated as:

A. n - 1

B. (r – 1) x (c – 1)

C. r x c - 1

55. A contingency table always involve:

A. Two degrees of freedom

B. Two dependent frequencies

C. Two independent variables

56. For any given level of significance, critical value of t:

A. Increases with increases in sample size

B. Increases with increases in degrees of freedom

C. Decreases with increases in degrees of freedom

57. For any given level of significance, critical value of chi-square:
A. Increases with increases in sample size
B. Increases with increases in degrees of freedom
C. Decreases with increases in degrees of freedom

58. The larger the discrepancy between the observed and expected frequencies for each cell in a contingency table:
A. the more likely the population proportions are the same
B. the more likely the null hypothesis will be rejected
C. the more likely the results will not be significant

59. The chi-square test requires that:
A. data be measured on a nominal scale
B. data conform to a normal distribution
C. expected frequencies are equal in all cells

60. Hypothesis testing involves:
A. deciding between two mutually exclusive hypotheses H_0 and H_1
B. deciding if the investigation was internally and externally valid
C. deciding if the differences between groups were large or small

Right answers

1-A; 2-A; 3-A; 4-C; 5-C; 6-C; 7-B; 8-A; 9-B; 10-B;

11-C; 12-B; 13-C; 14-C; 15-B; 16-A; 17-A; 18-A;

19-A; 20-A; 21-A; 22-A; 23-A; 24-B; 25-B; 26-B;

27-C; 28-C; 29-C; 30-C; 31-B; 32-C; 33-C; 34-B;

35-B; 36-C; 37-A; 38-A; 39-C; 40-A; 41-C; 42-A;

43-A; 44-A; 45-A; 46-B; 47-A; 48-A; 49-A; 50-B;

51-A; 52-A; 53-A; 54-B; 55-C; 56-C; 57-B; 58-B;

59-A; 60-A