Student T-Test for Comparison of Two Samples & Chi-square Test

Sorana D. Bolboacă

STATISTICS AND MEDICAL PRACTICE

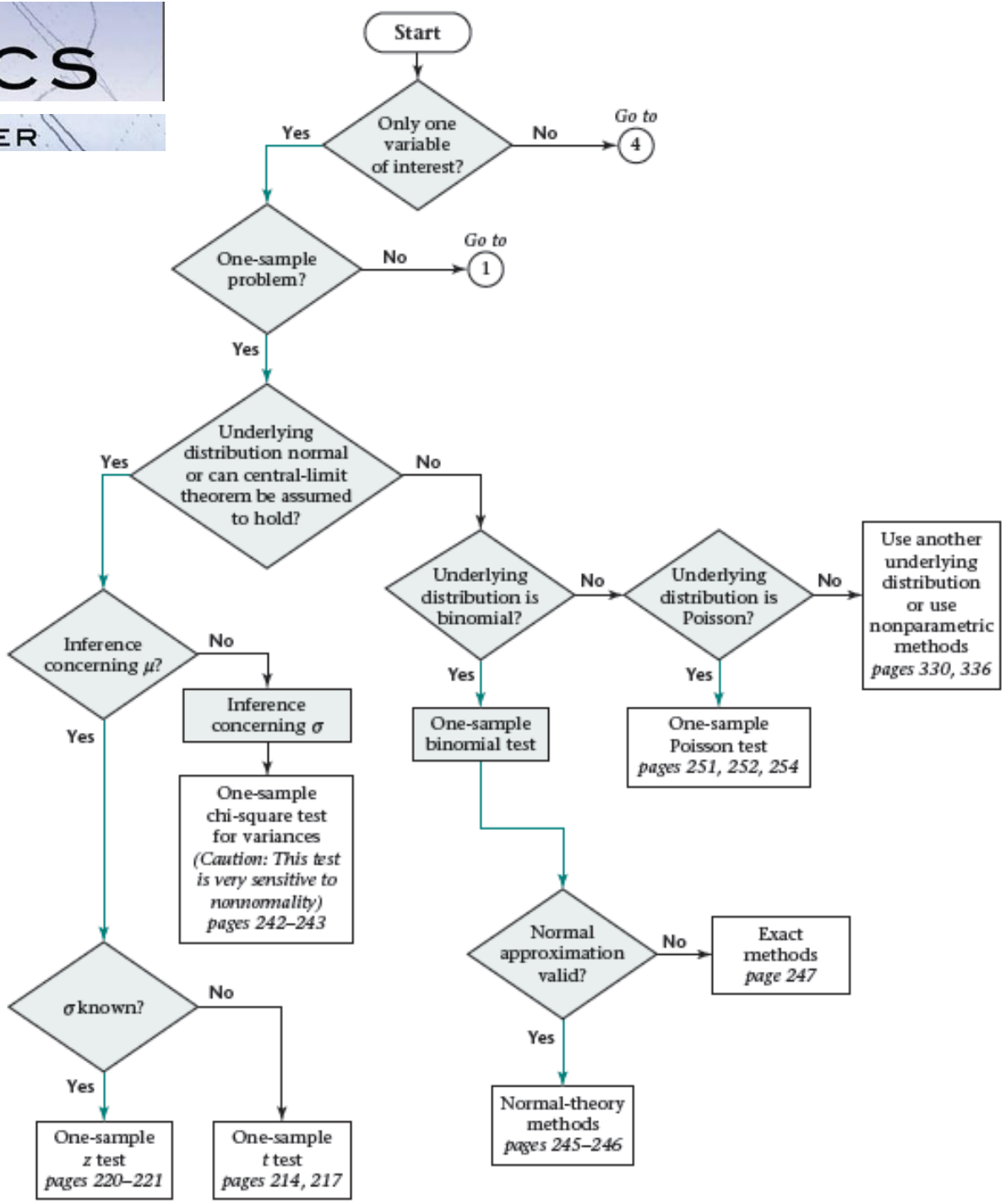
Source: http://kingeofdremes.wordpress.com/

OBJECTIVES

- Significance level versus p-value
- Student t-test for independent samples (parametric test)
- Chi-square test (non-parametric test)

SIGNIFICANCE LEVEL VS. p-VALUE

- significance level = property of a statistical procedure and takes a fixed value.
- p-value = random variable whose value depends upon the composition of the individual sample



STUDENT T-TEST FOR COMPARING TWO MEANS (UNKNOWN AND EQUAL VARIANCES)

- **Null hypothesis**: Means difference of the two populations is equal to zero.
- **Alternative hypothesis** for two-tailed test: Means difference of the two populations is NOT equal to zero.
- **Assumptions:**
  - The variables in the two samples are normal distributed
  - The variances are equal.
- If these two assumptions are not satisfied the test loss its validity.
- If the variances of populations are known the Z test is applied (is most powerful)

STUDENT T-TEST FOR COMPARING TWO MEANS (UNKNOWN AND EQUAL VARIANCES)

- **Degree of freedom (df):**
  - $df = n_1 + n_2 - 2$
- **Significance level**: $\alpha = 0.05$
- **Critical region** for two-tailed test

$$\left(-\infty;-t_{n_1+n_2-2;\frac{\alpha}{2}}\right] \cup \left[t_{n_1+n_2-2;\frac{\alpha}{2}};+\infty\right)$$

- **Statistics**

$$t = \frac{m_1 - m_2}{\sqrt{s\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$s = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$$

STUDENT T-TEST FOR COMPARING TWO MEANS (UNKNOWN AND EQUAL VARIANCES)

- We want to study whether there is a significant difference between the amount of blood uric acid in women from urban and rural. In a sample of 16 women aged between 30 and 50 years in urban areas, average uric acid was 5 mg/100 ml, with a variance of 2 mg/100 ml. An average equal to 4 mg/100 ml with a variance of 2 mg/100 ml was obtained on a sample of 16 women aged 30 to 50 years in rural areas.

STUDENT T-TEST FOR COMPARING TWO MEANS (UNKNOWN AND EQUAL VARIANCES)

- **Data**:
  - $n_1 = 16$; $n_2 = 16$
  - $m_1 = 5$; $m_2 = 4$
  - $s^2 = 2$
- **Null hypothesis**: There is no significant difference between the two samples means.
- **Alternative hypothesis** for two-sided test: There is a significant difference between the two samples means.

- **Degree of freedom**:

$df = n_1+n_2-2 = 16+16-2 = 30$

- **Significance level**:

$$\alpha = 0.05.$$

- **Critical region** for bilateral test:

$$(-\infty;-t_{n_1+n_2-2;0.025}] \cup [t_{n_1+n_2-2;0.025};+\infty)$$

$$(-\infty;-2.04] \cup [2.04;+\infty)$$

STUDENT T-TEST FOR COMPARING TWO MEANS (UNKNOWN AND EQUAL VARIANCES)

$$s = \sqrt{\frac{(n_1 - 1)\cdot s_1^2 + (n_2 - 1)\cdot s_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{(16-1)\cdot 2 + (16-1)\cdot 2}{16+16-2}} = \sqrt{\frac{60}{30}} = 1.41$$

$$t = \frac{m_1 - m_2}{\sqrt{s\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{5-4}{\sqrt{1.41\cdot\left(\frac{1}{16} + \frac{1}{16}\right)}} = \frac{1}{\sqrt{1.41\cdot 0.25}} = \frac{1}{\sqrt{0.3525}} = \frac{1}{0.5937} = 1.68$$

Conclusion:

- Statistical: The null hypothesis is failed to be rejected since the statistics did not belongs to the critical region.
- Clinical: The serum level of uric acid is not different in women from rural compared to those from urban areas.

of asphalt binder in low, intermediate and high service temperatures.

4.2.1. The mean hypothesis test method

Due to test conditions, the hypothesis test of two populations with unknown variances and equal number of samples is used in this paper. The steps of this method are [29]:

1. $H_0: \mu_1 - \mu_2 = 0$
2. $H_1: \mu_1 - \mu_2 > 0$ or $H_1: \mu_1 - \mu_2 < 0$
3. $T_{f,n_1+n_2-2} = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$

where H_0 is the null hypothesis, H_1 is the alternative hypothesis, T is the test statistics, $\overline{X}_1$ and $\overline{X}_2$ are the mean value measurements of two population samples, n_1 and n_2 are the numbers of each sample that has the equal quantity, and s_p^2 is the combined variance of the two samples which can be calculated as:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

where s_1^2 and s_1^2 are the variances of two samples.

4. T has the t-distribution with $n - 1$ degree of freedom, which α is the significant level.
5. The acceptable area is defined as:

$$T < t_{1-\alpha}$$

http://www.hindawi.com/journals/tswj/2013/608683/tab2/

Table 2: The values of blood pressures and heart rates prior to and after the treatment with metoprolol and nebivolol in the study

	Total (n = 60)	**Metoprolol** (n = 30)	**Nebivolol** (n = 30)	P **value**
Systolic blood pressure (mm/Hg)				
Before	152.1 ± 4.9	151.3 ± 3.6	152.8 ± 5.9	0.27**
After	136.4 ± 10.9	134.9 ± 10.5	137.8 ± 11.3	0.3**
P value	<0.001*	<0.001*	<0.001*	
Diastolic blood pressure (mm/Hg)				
Before	91.7 ± 3.9	91.1 ± 3.7	92.3 ± 4.1	0.25**
After	82.6 ± 7.1	82.4 ± 6.7	82.7 ± 7.3	0.87**
P value	<0.001*	<0.001*	<0.001*	
Heart rate (pulse/min)				
Before	75.7 ± 5.7	75.2 ± 5.7	76.2 ± 5.7	0.5**
After	70.2 ± 5.4	70.3 ± 5.8	70.1 ± 5.1	0.87**
P value	<0.001*	<0.001*	<0.001*	
Achieved targeted blood pressure (n, %)	37 (61.6)	19 (63.3)	18 (60)	0.5***

*Paired t-test; **independent samples t-test; ***chi-square test.

http://www.hindawi.com/journals/tswj/2013/608683/tab2/

Table 3: The values of blood pressures and heart rates measured during rest, exercise, and recovery periods in the study population.

	Total ($n = 60$)	**Metoprolol** ($n = 30$)	**Nebivolol** ($n = 30$)	P **value**
Systolic blood pressure (mm/Hg)				
Rest	137.4 ± 13.7	135.3 ± 13.5	139.6 ± 13.8	0.22
Stage 1	152 ± 22.5	151.7 ± 20.9	152.3 ± 24.4	0.91
Stage 2	162.7 ± 27.1	160.8 ± 27.1	164.6 ± 27.5	0.61
Stage 3	175.3 ± 29.3	174.3 ± 29.7	176.2 ± 29.7	0.84
Recovery (3 min)	164.4 ± 27.4	167.1 ± 29.8	161.7 ± 25	0.46
Recovery (5 min)	156.5 ± 27.9	158.5 ± 27	154.5 ± 29	0.58
Diastolic blood pressure (mm/Hg)				
Rest	82.6 ± 9.9	83.7 ± 10.6	81.5 ± 9.3	0.39
Stage 1	76 ± 21.2	78.6 ± 21.5	73.3 ± 20.9	0.34
Stage 2	72.9 ± 18.7	76.7 ± 16.5	68.9 ± 20.2	0.12
Stage 3	71.5 ± 19.5	77.3 ± 11.5	66.7 ± 23.4	0.07
Recovery (3 min)	77.9 ± 17	81.8 ± 16.6	74.1 ± 16.7	0.08
Recovery (5 min)	81.9 ± 12.9	83.2 ± 14.1	80.7 ± 11.8	0.45
Heart rate per minute				
Rest	71.7 ± 5.5	72.1 ± 5.3	71.2 ± 5.9	0.63
Stage 1	106.7 ± 16.4	105.2 ± 17.2	108.1 ± 15.7	0.49
Stage 2	122.8 ± 17.2	119.6 ± 16.6	126.3 ± 17.6	0.15
Stage 3	139.5 ± 18.5	140.3 ± 18.1	138.9 ± 19.2	0.82
Recovery (3 min)	118.7 ± 16.3	117.2 ± 17.7	120.1 ± 14.9	0.5
Recovery (5 min)	90.3 ± 14.5	89.4 ± 15.2	91.2 ± 14	0.63

Testing Association in Contingency Table

- We can perform a hypothesis test on a contingency table. The test we will use most often is the Chi-square test ($\chi 2$ test).
- χ^2 Test
  - Is proper to be applied if the sample size is large
  - The test is valid if the expected frequency of each cell is at least equal to 1 and the observed frequency is of 5
  - If the above-described conditions are not meet, the Fisher exact test is the proper test

χ^2 Test

- Indicate if that the two variables are or are not independent BUT DO NOT quantify the power of association between them.
- Steps:
  1. Define the hypotheses
  2. Define the parameter of the test
  3. Define the significance level
  4. Define the critical interval
  5. Calculate the observed value of the parameter of the test
  6. Make a decision

χ^2 TEST: PROBLEM

- The association between *Streptococcus mutans* (as risk factor) and dental caries was studied. A sample of 620 patients was investigated. The sample contains: 150 patients with caries and *Streptococcus mutans*, 230 patients without caries and without *Streptococcus mutans* and 60 patients with caries but without *Streptococcus mutans.* The presence of *Streptococcus mutans* is asscoiated with dental caries? (df=1; α=0.05; $\chi^2_{critical}$ = 3.84).

χ^2 Test: 1. Hypotheses

- H_0:
  - There is no association between *Streptococcus mutans* and dental caries.
  - The presence of *Streptococcus mutans* and dental caries are independent.
- H_1/H_a:
  - There is an association between *Streptococcus mutans* and dental caries.
  - The presence of *Streptococcus mutans* and dental caries are not independent.

χ^2 Test: 2. Parameter of the test

$$\chi^2 = \sum_{i=1}^{r \cdot c} \frac{(f_i^0 - f_i^t)^2}{f_i^t}$$

Follow a distribution law with (r-1)·(c-1) degree of freedom

where

- χ^2 = the parameter of χ^2 test
- f_i^o = observed frequency
- f_i^t = expected/theoretic frequency

χ^2 Test: 3. Significance level

- Let $\alpha = 0.05$ (5%) be the significance level.

χ^2 Test: 4. Critical region

- Critical region: $[\chi_\alpha{}^2, \infty)$
- For $\alpha = 0.05$:
  - $\chi_\alpha{}^2 = 3.84$
  - $[3.48, \infty)$

χ^2 TEST: 5. PARAMETER OF THE TEST

observed	DC+	DC-	Total
SP +	TP = **150**	FP = **180**	330
SP -	FN = **60**	TN = **230**	290
Total	210	410	620

expected	DC+	DC-	Total
SP +	**= 330×210/620**	**= 330×410/620**	330
SP -	**= 290×210/620**	**= 290×410/620**	290
Total	210	410	620

χ^2 Test: 5. Parameter of the test

observed	DC+	DC-
SP +	150	180
SP -	60	230

expected	DC+	DC-
SP +	= 112	= 218
SP -	= 98	= 192

$$\chi^2 = \frac{(150-112)^2}{112} + \frac{(180-218)^2}{218} + \frac{(60-98)^2}{98} + \frac{(230-192)^2}{192}$$

$$\chi^2 = \frac{38^2}{112} + \frac{(-38)^2}{218} + \frac{(-38)^2}{98} + \frac{(38)^2}{192}$$

$$\chi^2 = \frac{1444}{112} + \frac{1444}{218} + \frac{1444}{98} + \frac{1444}{192} = 12.89 + 6.63 + 14.73 + 7.52 = 41.77$$

x^2 Test: 6. Making decision

- If $\chi^2 \in [3.84, \infty)$ H_0 is rejected with a risk of error of type I (α).
- If $\chi^2 \notin [3.84, \infty)$ H_0 is accepted with a risk of error of type II (β).

- Since $41.77 \in [3.84, \infty)$ H_0 is rejected with a risk of error of 5%.
- **There is an association between *Streptococcus mutans* and dental caries.**

CONTINUITY CORRECTION (YATES'S CORRECTION)

- For small sample sizes the χ^2 test is too likely to reject the null hypothesis (it tends to spot differences where none really exist).
  - A continuity correction can be made to allow for this.
  - Two conditions have to be met:
    - All expected frequencies must be greater than 1
    - 80% of observed frequencies must be greater than 5

Yates's correction

$$\chi^2 = \sum_{i=1}^{r \cdot c} \frac{|f_i^0 - f_i^t|^2 - 0.5}{f_i^t}$$

FISHER'S EXACT TEST

- Chi-square procedures can be legitimately applied only if all values of **E** are equal to or greater than 5.
- If a 2×2 contingency table fails to meet the conditions required for the χ^2 test then Fisher's exact test can be used.
- It is based on different mathematics to the χ^2 test which are more robust when sample sizes are small.

FISHER'S EXACT TEST

- H_0: there is no association between smoking and dental caries
- If the null hypothesis is true - if any ostensible association between smoking and dental caries were the result of nothing more than mere chance coincidence -how likely is it that we might end up with a result this large or larger?

observed	DC+	DC-	Total
smoking +	TP = 2	FP = 7	9
smoking -	FN = 8	TN = 2	10
Total	10	9	19

FISHER'S EXACT TEST

- Suppose that the initial assessment was performed and the number of subjects who do and do not show characteristics (smoking and dental caries) were counted, but have not yet sorted the subjects according to the correspondences of smoking and dental caries. In this case, all they would have would be the marginal totals shown in the following table/
- Given these marginal totals, there are 10 possible ways in which the specific correspondences between smoking and dental caries.

	DC+	DC-	Total
smoking +			9
smoking -			10
Total	10	9	19

FISHER'S EXACT TEST

- The p-value is calculated directly from the formula:

$$p = \frac{(a+c)!(b+d)!(c+d)!(a+b)!}{n!a!b!c!d!}$$

- The p-value for the observed contingency table must be added to the p-value of the more extreme contingency table.

FISHER'S EXACT TEST

Obs	DC+	DC-	Total
smoking +	6	2	8
smoking -	1	6	7
Total	7	8	15

Exp	DC+	DC-	Total
smoking +	7	1	8
smoking -	0	7	7
Total	7	8	15

- The p-value must be calculated for the two contingency tables:

$$p_1 = \frac{7!8!7!8!}{15!6!2!6!} = 0.0305 \qquad p_2 = \frac{7!8!7!8!}{15!7!0!7!} = 0.0012$$

- Therefore $p = p_1 + p_2 = 0.0305 + 0.0012 = 0.0317$

- The p-value = $0.0317 < \alpha = 0.05 \Rightarrow$ that smoking is associated with dental caries.

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3663126/pdf/WJC-5-124.pdf

AIM: To determine whether there are gender differences in the epidemiological profile of atrial fibrillation (AF) and to characterise the clinical, biochemical, and therapeutic factors associated with AF.

Gender differences

Table 2 shows the differences between the men and the women in epidemiological, biochemical, BP and therapeutic characteristics. The women exhibited more obe-

Table 2 Epidemiological, clinical and therapeutic differences between genders (mean ± SD) *n* (%)

	Females (*n* = 542)	Males (*n* = 486)	*P* value
Mean age, yr	72.7 ± 5.8	72.8 ± 5.8	NS
Abdominal circumference, cm	96.6 ± 11.8	100.4 ± 11.0	< 0.001
Weight, kg	71.4 ± 11.5	79.5 ± 11.5	< 0.001
Mean height, cm	155.2 ± 6.7	166.7 ± 6.7	< 0.001
BMI	29.6 ± 4.5	28.6 ± 3.6	< 0.001
Obesity	224 (41.4)	160 (32.9)	0.005
Years from the onset of HT	11.0 ± 8.2	10.8 ± 8.1	NS
Diabetes mellitus	134 (24.7)	150 (30.9)	0.03
Dyslipidaemia	267 (49.3)	230 (47.3)	NS
Smokers	17 (3.1)	76 (15.6)	< 0.001
Sedentariness	352 (70.5)	274 (56.4)	< 0.001
Regular alcohol intake	5 (0.9)	33 (6.8)	< 0.001

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3663126/pdf/WJC-5-124.pdf

AIM: To determine whether there are gender differences in the epidemiological profile of atrial fibrillation (AF) and to characterise the clinical, biochemical, and therapeutic factors associated with AF.

Table 4 Treatment differences between genders in patients with atrial fibrillation (n = 106) n (%)

	Females (n = 50)	Males (n = 56)	P value
Diuretics	34 (68.0)	30 (53.6)	NS
Beta-blockers	16 (32)	17 (30.4)	NS
Calcium antagonists	15 (30)	5 (8.9)	0.007
ACEI	12 (24)	14 (25)	NS
ARB	32 (64)	31 (55.4)	NS
Antiplatelet agents	4 (8)	16 (26.8)	0.010
VKA	29 (58)	25 (44.6)	NS
ATG or VKA	33 (66)	41 (71.4)	NS

ACEI: Angiotensin-converting enzyme inhibitors; ARB: Angiotensin receptor blocker; VKA: Vitamin K antagonist; ATG: Anti-aggregants. NS: No significant.