Tests on Frequencies

Sorana D. Bolboacă

OBJECTIVES

- Contingency tables
- Chi-square test
- Fisher exact test
- Tests on proportions

2×2 Contingency Tables

- Nominal scale: dichotomiale: 2-by-2 contingency Table)
- Ordinal scale: r-by-c contingency table
- Absolute frequency (number of events per category)
- 2×2 contingency table: 4 categories
  - TP = true pozitive
  - FP = false pozitive
  - FN = false negative
  - TN = true negative

2×2 CONTINGENCY TABLES

	Caries +	Caries -	Total
Fluoridated +	TP = 77	FP = 29	= 77+29 = 106
Non-Fluoridated -	FN = 95	TN = 31	= 95+31 = 126
Total	=77+95=172	=29+31=60	= 77+29+95+31 =232

- Degree of freedom (df) = the minimum number of values in cells necessary to compute the values from other cells
  - 2×2 contingency table: if we have the totals on rows and column the values in all 4 cells could be computed
  - df = (r - 1)(c - 1); r = number of rows, c = number of columns

RISK

- Risk means the same thing as probability to a mathematician. Clinicians and epidemiologists tend to use the word risk in a particular way but we still calculate risks in the same way as any other probability.
- The risk of an outcome is the number of times the outcome of interest occurs divided by the total number of possible outcomes.

RISK

- In the paper *Caries prevalence in northern Scotland before and 5 years after, water defluoridation* (Stephen et al., 1987, BDJ 163: 324-326) the researchers studied two groups of children in Wick; one group whilst the water was fluoridated and one group after defluoridation. Out of 106 children examined whilst the water was fluoridated 77 had caries.
- We get the risk of caries whilst the water was fluridated by the following calculation:

Risk = 77 / 106 = 0.73

Risk Ratio (Relative risk)

	Caries +	Caries -	Total
Fluoridated +	TP = 77	FP = 29	= 77+29 = 106
Non-Fluoridated -	FN = 95	TN = 31	= 95+31 = 126
Total	=77+95=172	=29+31=60	= 77+29+95+31 =232

- If we want to compare the effects of fluoridated and non-fluoridated water we could calculate the risk of having caries for each group:
- Risk of having caries when water is fluoridated = 77 / 106 = 0.73
- Risk of having caries when water is not fluoridated = 95 / 126 = 0.75

RISKS COMPARISON

- We can compare the risk for each of the groups using the risk ratio. The risk ratio for being caries free when water is fluoridated compared to when it is not fluoridated is:
- **(Risk when fluoridated) / (Risk when not fluoridated) = 0.73 / 0.75 = 0.96**
- The risk of having caries when the water is fluoridated is only 0.96 that of when the water is not fluoridated.
- An risk ratio of 1 means there is no difference between the groups
- The 95% CI includes 1 so we have a (statistically) non-significant result.

ODDS

	Caries +	Caries -	Total
Fluoridated +	TP = 77	FP = 29	= 77+29 = 106
Non-Fluoridated -	FN = 95	TN = 31	= 95+31 = 126
Total	=77+95=172	=29+31=60	= 77+29+95+31 =232

- The odds in favor of a particular outcome is the number of times the outcome occurs divided by the number of times it does not occur.
- 77 children who had caries and 29 who didn't:

Odds = 77 / 29 = 2.66

ODDS

- Odds of less than 1 mean the outcome occurs less than half the time
- Odds of 1 mean the outcome occurs half the time
- Odds of more than 1 mean the outcome occurs more than half the time

ODDS RATIO

- If we want to compare the effects of fluoridated and non-fluoridated water we could calculate the odds for each group:
  - Odds for having caries when water is fluoridated = 77 / 29 = 2.66
  - Odds for having caries when water is not fluoridated = 95 / 31 = 3.06

ODDS RATIO

- We can compare the odds using the odds ratio. The odds ratio for having caries when water is fluoridated compared to when it is not fluoridated is:

(Odds when fluoridated) ÷ (Odds when not fluoridated) = 2.66 / 3.06 = 0.87

- So, the odds of having caries when the water is fluoridated are about 90% those of when the water is not fluoridated.
- An odds ratio of 1 means there is no difference between the groups

Risks and Odds: Other Measures of Association

Denumire	Formula	Definiţie
False positive rate	=FP/(FP+TP)	Probability of a false positive test (α)
False negative rate	=FN/(FN+TP)	Probability of a false negative test (β)
Sensibility	=TP/(TP+FN)	Probability of a true positive test (1- β)
Specificity	=TN/(TN+FP)	Probability of a true negative test (1- α)
Accuracy	=(TP+TN)/n	General probability of a correct decision
Positive predictive value	=TP/(TP+FP)	Probability of a correct positive test
Negative predictive value	=TN/(TN+FN)	Probability of a correct negative test
Relative risk	=[TP(FP+TN)]/[FN(TP+FP)]	
Odd ratio	=(TP·TN)/(FN·FP)	
Attributable risk	=TP/(TP+FP)-FN/(FN+TN)	

RISKS AND ODDS: MEASURES OF ASSOCIATION

	Caries +	Caries -	Total
Fluoridated +	**TP = 77**	**FP = 29**	**= 77+29 = 106**
Non-Fluoridated -	**FN = 95**	**TN = 31**	**= 95+31 = 126**
Total	**=77+95=172**	**=29+31=60**	**= 77+29+95+31 =232**

Name	Formula
False positive rate	= 29/(77+29) = 0.2736
False negative rate	= 95/(95+31) = 0.7540
Sensibility	= 77/(77+95) = 0.4477
Specificity	= 31/(31+29) = 0.4833
Accuracy	= (77+31)/232 = 0.4655
Positive predictive value	= 77/(77+29) = 0.7264
Negative predictive value	= 31/(31+95) = 0.2460
Relative risk	= 77(29+31)/95(77+29) = 0.4588
Odd ratio	= (77·31)/(95·29) = 0.8664
Attributable risk	= 77/(77+29)-95/(95+31) = -0.0275

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 (χ2 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 at least 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 statistical 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**

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

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

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

χ^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 fail to be rejected 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 = $0.0317 < \alpha = 0.05 \Rightarrow$ smoking is associated with dental caries.

Tests of Frequencies

- Comparing an observed frequency with a theoretical frequency
- Testing the equality of two frequencies

Z Test: 1. Comparing an observed frequency with a theoretical frequency

- **Aim:** Investigation of the significance of the difference between a theoretical frequency *p* (observed in a population) and an observed frequency *f* (observed on a representative sample) (qualitative variable).
- **Assumptions*:***
  - The test is correctly applied when sample size is large enough ($n \cdot p > 10$ & $n \cdot (1-p) > 10$, where n = sample size; p = frequency).
- **Statistic**:

$$z = \frac{f - p}{\sqrt{\frac{p(1-p)}{n}}}$$

  - n = sample size

Z TEST: 1. COMPARING AN OBSERVED FREQUENCY WITH A THEORETICAL FREQUENCY

We are interested in investigating the frequency of hepatitis B in personnel working in laboratories at Infectious Disease Clinics in Transylvania. It is known from previous studies that the prevalence of hepatitis B in the general population of Transylvania is 9%. In a studied sample of 100 people a frequency of hepatitis B of 6% was obtained. There is significant difference between the frequency of hepatitis B in personnel working in hospitals for infectious diseases laboratory in Transylvania from the general population?

- $f = 0.06$; $p = 0.09$; $n = 100$
- H_0: No significant difference in frequency of hepatitis B in the sample studied from the frequency of hepatitis B in the general population.
- H_a/H_1, two-tailed test: There is a significant differences in the sample frequency of hepatitis B and hepatitis B prevalence in the general population.

Z TEST: 1. COMPARING AN OBSERVED FREQUENCY WITH A THEORETICAL FREQUENCY

- f = 0.06; p = 0.09; n = 100
- Significance level (type I error): α = 0.05
- Critical region, two-tailed test:
  - (-∞; -1.96]∪[1.96; ∞)

$$z = \frac{f - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.06 - 0.09}{\sqrt{\frac{0.09(1-0.09)}{100}}} = \frac{-0.03}{\sqrt{\frac{0.09 \cdot 0.91}{100}}}$$

$$z = \frac{-0.03}{\sqrt{\frac{0.0819}{100}}} = \frac{-0.03}{\sqrt{0.000819}} = \frac{-0.03}{0.029} = -1.05$$

Z TEST: 1. COMPARING AN OBSERVED FREQUENCY WITH A THEORETICAL FREQUENCY

- **Conclusions**:
  - **Statistical**: Since the calculated statistic of the test did not belong of the critical region, the null hypothesis fails to be rejected.
  - **Clinical**: No significant difference in frequency of hepatitis B in the studied sample from the frequency of hepatitis B in the general population was observed.

Z TEST: 1. TESTING THE EQUALITY OF TWO FREQUENCIES

- **Aim**: Investigation of the significance of the difference between relative frequencies
  - Qualitative variable on two independent random samples taken from two different populations.
- **Assumptions:**
  - The number of observations in each sample is large enough ($n_1 > 30$ AND $n_2 > 30$)

$$z = \frac{(p_1 - p_2)}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$p = \frac{p_1 n_1 + p_2 n_2}{n_1 + n_2}$$

Z TEST: 1. TESTING THE EQUALITY OF TWO FREQUENCIES

- HIV status was studied in a sample of 170 women aged between 18 and 40 years in Moldova, and in a sample of 89 women (same range of aged) in Transylvania. The frequency of HIV+ was of 10% in Moldova and of 2.7% in Transylvania.
- The frequency of HIV infection in women between 18 and 40 years in Moldova is different from the frequency of infection in women of the same age in Transylvania?

Research data: $p_1 = 0.10$; $p_2 = 0.027$; $n_1 = 170$; $n_2 = 89$

H_0: The frequency of HIV infection in women from Moldova is not significantly different by the frequency of HIV infection in women from Transylvania.

H_a/H_1, two-sided test: The frequency of HIV infection in women from Moldova is significantly different by the frequency of HIV infection in women from Transylvania

Z TEST: 2. TESTING THE EQUALITY OF TWO FREQUENCIES

$$z = \frac{(p_1 - p_1)}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.10 - 0.027}{\sqrt{0.075 \cdot (1 - 0.075)\left(\frac{1}{170} + \frac{1}{89}\right)}}$$

$$z = \frac{0.073}{\sqrt{0.075 \cdot 0.925 \cdot (0.006 + 0.011)}} = \frac{0.073}{\sqrt{0.001}} = \frac{0.073}{0.034} = 2.118$$

Significance level: $\alpha = 0.05$

Critical region:

- Two-sided test: $(-\infty; -1.96\,] \cup [1.96; \infty)$

Conclusion:

- ***Statistical***: The null hypothesis is rejected since the calculated statistic belongs to the critical region.
- ***Clinical***: The frequency of HIV infection in women from Moldova is significantly different compared to the frequency of HIV infection in women from Transylvania.

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.

TESTS ON FREQUENCIES – PROBLEM

The set of women with at least one birth was arbitrarily divided into two categories: (1) women whose age at first birth was ≤29 years and (2) women whose age at first birth was ≥30 years. The following results were found among women with at least one birth: 683 of 3220 (21.2%) women with breast cancer (case women) and 1498 of 10,245 (14.6%) women without breast cancer (control women) had an age at first birth ≥30. How can we assess whether this difference is significant or simply due to chance?

- p_1 = the probability that age at first birth is ≥30 in case women with at least one birth (=683/3220 = 0.212) and p_2 = the probability that age at first birth is ≥30 in control women with at least one birth (=1498/10245 = 0.146).
- H_0: $p_1 = p_2 = p$ vs. H_1: $p_1 \neq p_2$ for some constant p.
- p = (683 + 1498)/(3220 + 10,245) = 2181/13,465 = 0.162 ⟶ q = 1 – 0.162 = 0.838
- Since n_1pq = 3220·(0.162) ·(0.838) = 437 ≥ 5 and n_2pq = 10245·(0.162)·(0.838) = 1391 ≥ 5 ⟶ z test is proper to be use

TESTS ON FREQUENCIES

The set of women with at least one birth was arbitrarily divided into two categories: (1) women whose age at first birth was ≤29 years and (2) women whose age at first birth was ≥30 years. The following results were found among women with at least one birth: 683 of 3220 (21.2%) women with breast cancer (case women) and 1498 of 10,245 (14.6%) women without breast cancer (control women) had an age at first birth ≥30. How can we assess whether this difference is significant or simply due to chance?

- $z = \dfrac{|0.212-0.146|-\left(\frac{1}{2\cdot 3220}+\frac{1}{2\cdot 10245}\right)}{\sqrt{0.162\cdot 0.838\cdot\left(\frac{1}{3220}+\frac{1}{10245}\right)}} = \dfrac{0.0657}{0.00744}$, z = 8.8
- p = 1.53E-18
- Therefore, we can conclude that women with breast cancer are significantly more likely to have had their first child after age 30 than are comparable women without breast cancer.

RECALL

Conditions for the $\chi 2$ test

All expected values must be greater than 1

80% of expected values must be greater than 5

Online calculator

r×n: Contingency Table

Online calculator