BAU
BAHÇEŞEHİR ÜNİVERSİTESİ

BME 3005
Biostatistics

Lecture 6: *Rates and Proportions*

Burcu Tunç Çamlıbel

Introduction

- So far, we have learned
  - how to summarize the data
    - Mean, variance, std, median, percentiles
    - Standard error of the mean to estimate the precision with which a sample mean estimates the population mean
  - ANOVA
    - If the hypothesis that all the samples were drawn from a single population is true, then the within group or between group or the real population variances should be almost EQUAL & F should be close to 1.

$$F = \frac{\text{population variance estimated from sample means}}{\text{population variance estimated as average of sample variances}}$$

$$F = \frac{s^2_{bet}}{s^2_{wit}}$$

Introduction

- t-test - to analyze the differences among two groups consisting of interval data

$$t = \frac{\text{difference in sample means}}{\text{standard error of difference of sample means}} \qquad t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{(s_1^2/n) + (s_2^2/n)}}$$

- Multiple comparison tests to correct for using t-test for multiple pairwise comparisons
  - Bonferroni t-test -- cutoff p level = α_T/k
  - Holm test - For the j^{th} hypothesis, use $\alpha_j = \alpha_T/(k-j+1)$ and j=1…k
  - SNK test - check Table 4.3 with $v_d = m(n-1)$ degrees of freedom and p which is the difference between the number of means being tested

$$q = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{s_{wit}^2}{2}\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}}$$

Introduction

- Tukey test – same as SNK, the only difference is the value p is set to # of groups in the study.
- Holm test is recommended for most of the multiple comparison cases.
- Multiple comparisons against a single control group
  – 2nd Bonferroni or 2nd Holm tests - The t test statistic is computed as before, and the number of comparisons is reduced to reflect the total number of comparisons made.
  – Dunnett's test - Analog of SNK for multiple comparisons against a single control group

$$q' = \frac{\bar{X}_{\text{con}} - \bar{X}_A}{\sqrt{s^2_{\text{wit}}\left(\frac{1}{n_{\text{con}}} + \frac{1}{n_A}\right)}}$$

Summary of Statistical Methods (table at the cover)

Summary of Some Statistical Methods to Test Hypotheses

Scale of measurement	Type of experiment: Two treatment groups consisting of different individuals	Three or more treatment groups consisting of different individuals	Before and after a single treatment in the same individuals	Multiple treatments in the same individuals	Association between two variables
Interval (and drawn from normally distributed populations*)	Unpaired *t* test (Chapter 4)	Analysis of variance (Chapter 3)	Paired *t* test (Chapter 9)	Repeated-measures analysis of variance (Chapter 9)	Linear regression, Pearson product-moment correlation, or Bland-Altman analysis (Chapter 8)
Nominal	Chi-square analysis-of-contingency table (Chapter 5)	Chi-square analysis-of-contingency table (Chapter 5)	McNemar's test (Chapter 9)	Cochrane Q†	Relative rank or odds ratio (Chapter 5)
Ordinal‡	Mann-Whitney rank-sum test (Chapter 10)	Kruskal-Wallis statistic (Chapter 10)	Wilcoxon signed-rank test (Chapter 10)	Friedman statistic (Chapter 10)	Spearman rank correlation (Chapter 8)
Survival time	Log-rank test or Gehan's test (Chapter 11)				

*If the assumption of normally distributed populations is not met, rank the observations and use the methods for data measured on an ordinal scale.

†Not covered in this text.

‡Or interval data that are not necessarily normally distributed.

BAU
BAHÇEŞEHİR ÜNİVERSİTESİ

Nominal data

- It does not get measured on a scale with constant intervals.
  – example interval data: cm, kg, days etc.
- In nominal data there is no arithmetic relationship between the different classifications.
- Example nominal data
  – male, female
  – Caucasian, African American, Asian
  – Dead or alive

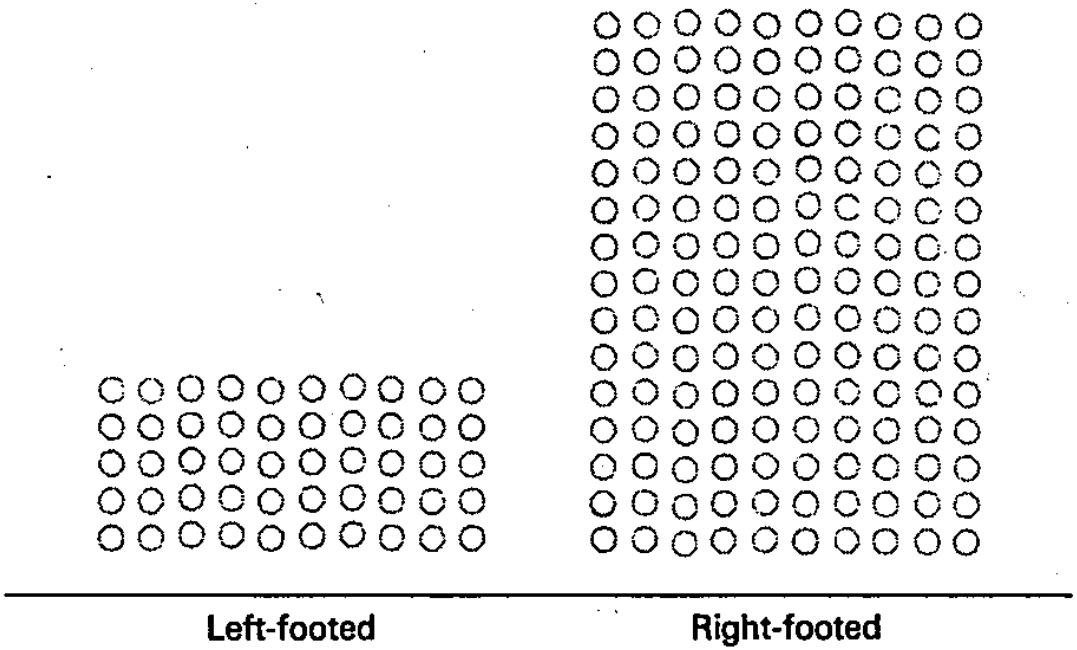
BAU
BAHÇEŞEHİR ÜNİVERSİTESİ

How to Describe Nominal Data

- Simple solution: Count the number of occurrences and compute the corresponding percentages

- How can we estimate the precision with which percentages based on limited samples approximate the true rates that would be observed if we were to examine the entire population?

Figure 5.1



50 left footed
150 right footed

p_{left}= 50/200 = 0.25
p_{right} = 150/200 = 0.75
$p_{left}+p_{right}$ = 1

When there are only 2 possible classes and they are mutually exclusive we can describe the whole population with a single p value. The other proportion will be 1-p.

p

- p is the proportion of members with one of the attributes.
- p is the probability of drawing a left-footed Martian at random in our scenario.
- X=1 for each left footed Martian, X=0 for right footed.

$$\mu = \frac{\Sigma X}{N} = \frac{1 + 1 + \cdots + 1 + 0 + 0 + \cdots + 0}{200}$$

$$= \frac{50(1) + 150(0)}{200} = \frac{50}{200} = 0.25$$

$\mu = p_{left}$

Mean for nominal data

- Suppose M members of a population with N individuals have some attribute and N-M members do not.

- Let's assign X=1 for the members having the attribute, X=0 for others

$$\mu = \frac{\Sigma X}{N} = \frac{M(1) + (N - M)(0)}{N} = \frac{M}{N} = p$$

Standard deviation of a nominal data

- Even though there are only two possible values (0,1) in the case of left/right footedness, the standard deviation will differ based on the p value.

$$\sigma = \sqrt{\frac{\Sigma(X - \mu)^2}{N}}$$

BAU
BAHÇEŞEHİR ÜNİVERSİTESİ

Standard deviation of a nominal data

- Remember, μ=p, M members have a value of 1, N-M are 0.

$$\sigma = \sqrt{\frac{(1-p)^2+(1-p)^2+\cdots+(1-p)^2+(0-p)^2+(0-p)^2+\cdots+(0-p)^2}{N}}$$

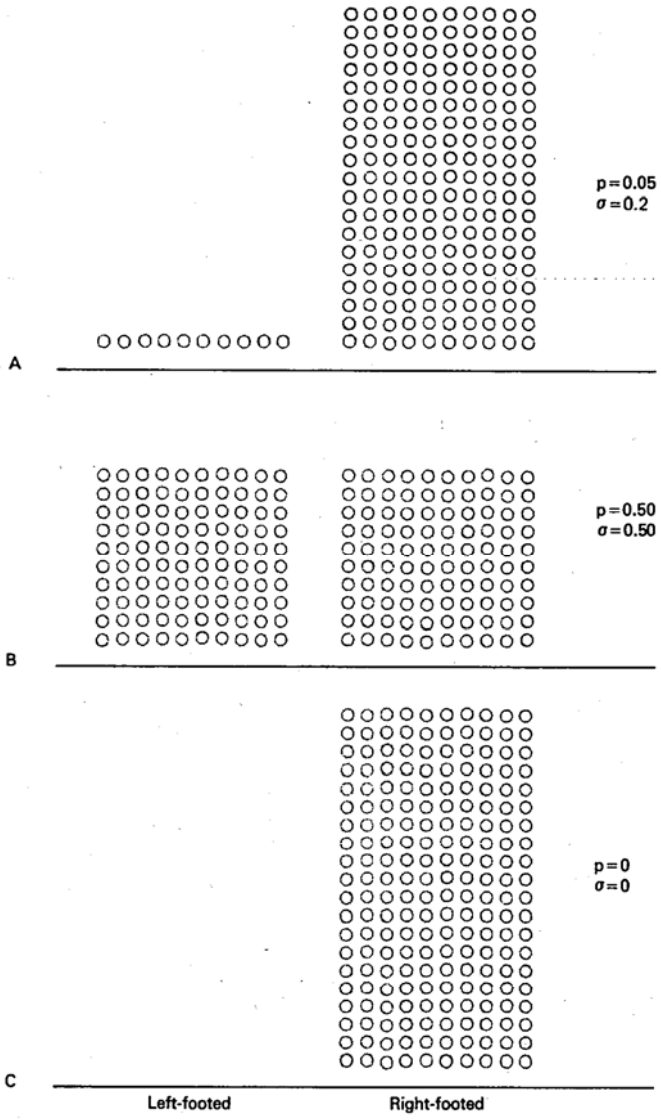
$$= \sqrt{\frac{M(1-p)^2+(N-M)p^2}{N}} = \sqrt{\frac{M}{N}(1-p)^2+\left(1-\frac{M}{N}\right)p^2}$$

- Remember M/N=p

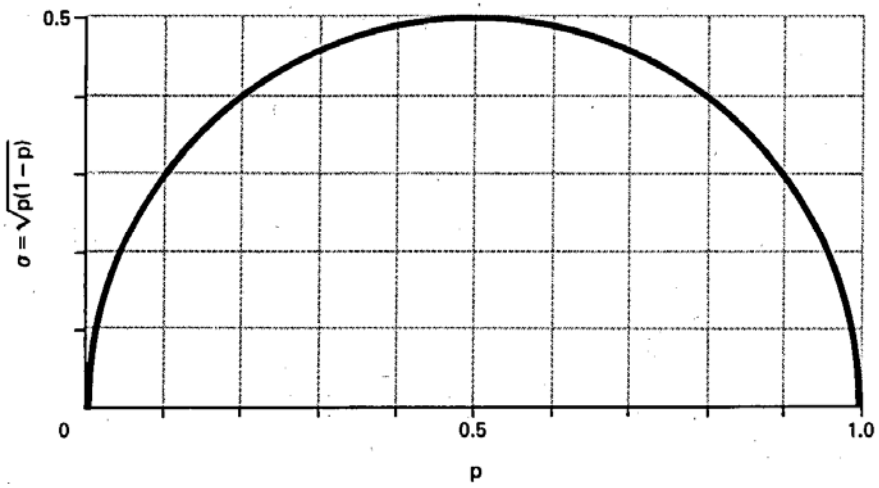
$$\sigma = \sqrt{p(1-p)^2+(1-p)p^2} = \sqrt{[p(1-p)+p^2](1-p)}$$

$$\sigma = \sqrt{p(1-p)}$$

Examples



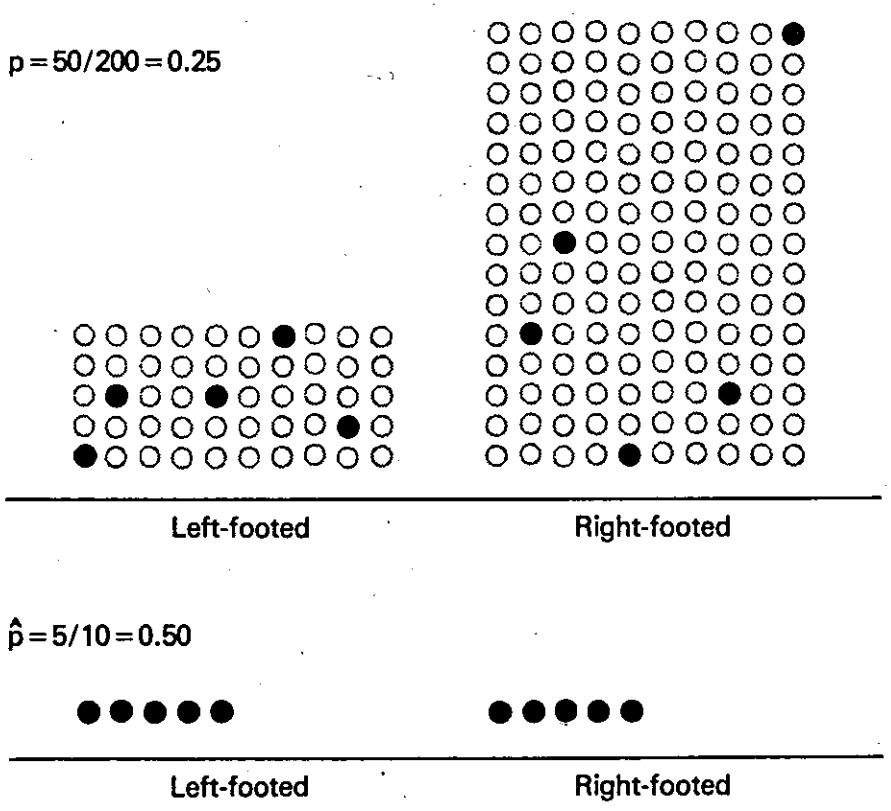
p and σ relation



- maximum standard deviation when p=0.5 (any given member is as likely to have the attribute as not
- σ=0 when p=0 or p=1 (no variability, everybody has a feature or not)
- Σ only depends on p and does not provide additional information like the mean and the std of a normally distributed variable

Estimating proportions from samples

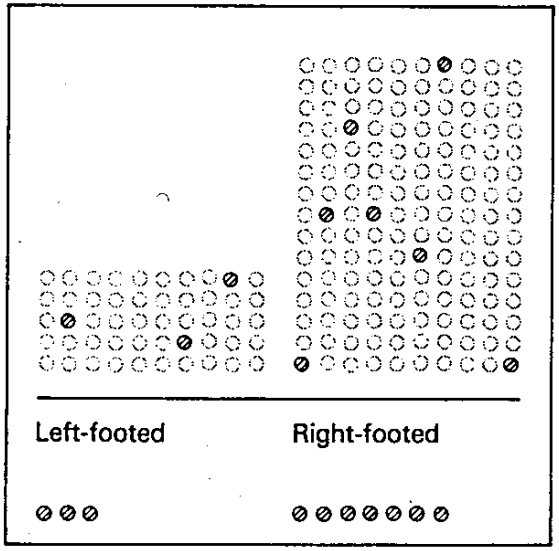
How well does the proportion estimated from a sample represent the whole population?

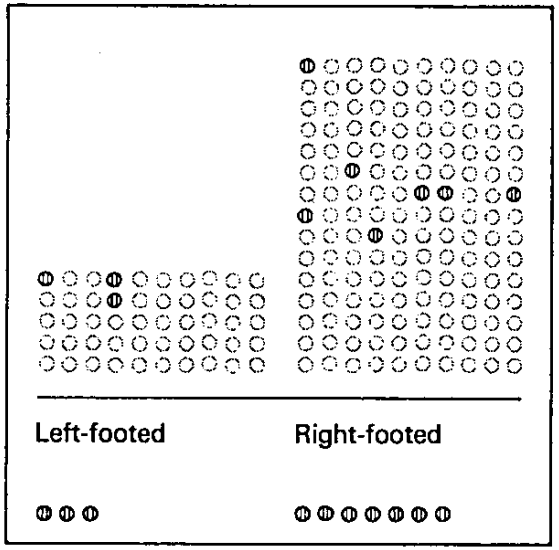


p=0.25 for the whole population

p^ = 0.50 for this sample

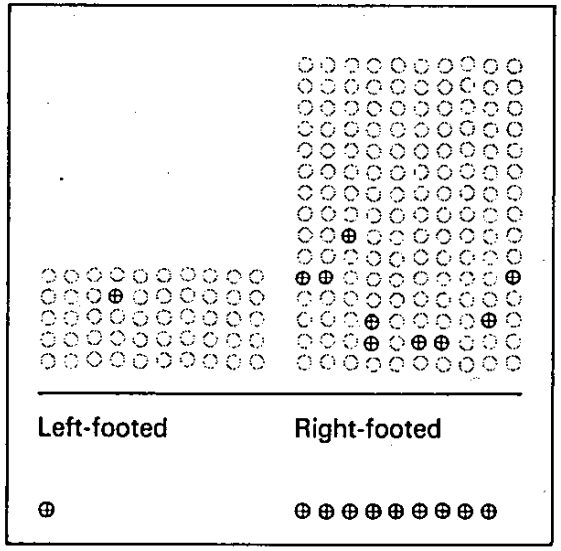
p^= 0.3

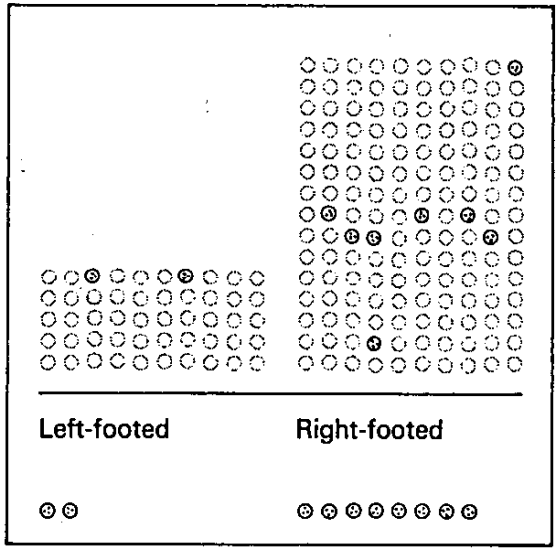




p^= 0.3

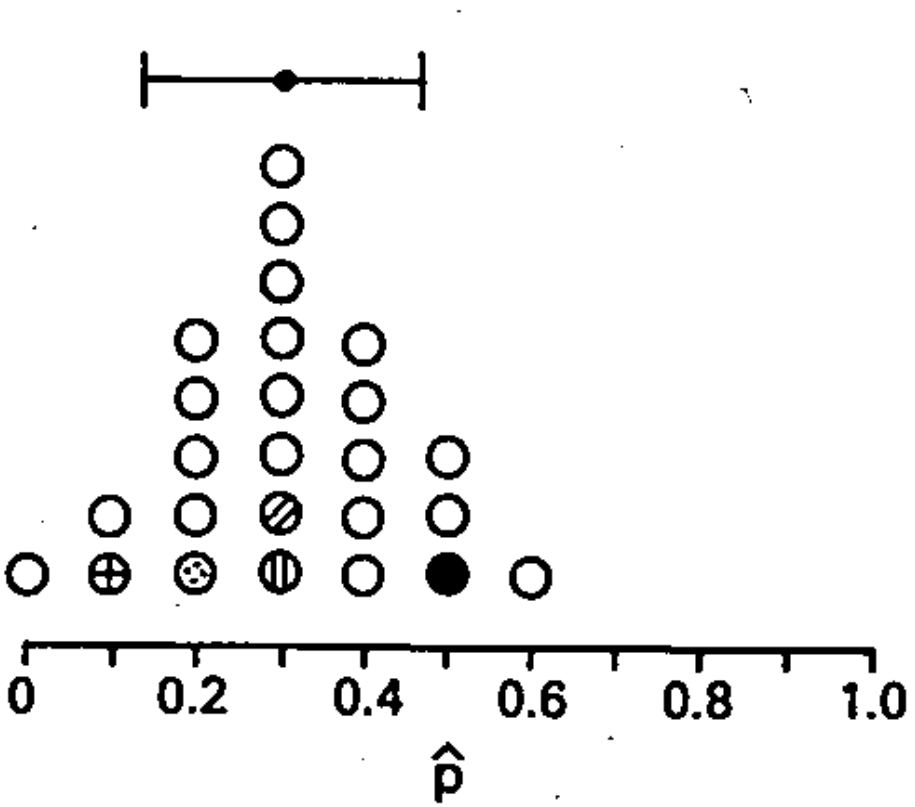
p^= 0.1





p^= 0.2

The standard error of estimate of the proportion ($\sigma_{p\hat{}}$)



- There will be a distribution to the p^ values.
- 25 sample p^ values
- The standard deviation of this curve is the standard error of the estimate of the proportion. (0.14 for this case)
- Just like SEM, $\sigma_{\hat{p}} = \frac{\sigma}{\sqrt{n}}$

SE of proportion

$$\sigma = \sqrt{p(1-p)}$$

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

Standard error of the proportion for sample with p^

$$s_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

BAU
BAHÇEŞEHİR ÜNİVERSİTESİ

Central limit theorem

- The distribution of sample means will be approximately normal regardless of the distribution of the values in the original population from which the samples were drawn.

- The mean value of the collection of all possible sample means will be equal to the mean of the original population.

- The standard deviation of the collection of all possible means of samples of a given size, called the standard error of the mean, depends on both the standard deviation of the original population and the size of the sample.

BAU
BAHÇEŞEHİR ÜNİVERSİTESİ

SE of the proportion

- SE of the proportion, p^, will be normally distributed just as the CLT states.
- The mean of the p^ distribution will be equal top, with an std of $\sigma_{p^}$.
- On the other hand, this assumption fails for p near 0 or 1, or when the sample size is too small. np^ and n(1-p^) both should exceed about 5 for this approximation to work.
- About 95% of a normally distributed population lies within 2*std of the mean. So, 95% of p^ will lie within 2*$s_{p^}$ of p.

BAU
BAHÇEŞEHİR ÜNİVERSİTESİ

Independent Bernoulli trials

- *Each individual trial has two mutually exclusive outcomes.*
- *The probability p of a given outcome remains constant.*
- *All the trials are independent.*

- In terms of a population:
  – Each member of the population belongs to one of the two classes.
  – The proportion of members of the population in one of the classes p remains constant.
  – Each member of the sample is selected independently of all other members.

BAU
BAHÇEŞEHİR ÜNİVERSİTESİ

z test

- p^ is analogous to the sample mean, and we also have developed a standard error of p^.

- z-test is analogous to t-test and used to test the hypothesis that two samples were drawn from populations containing the same proportion of individuals with a given attribute.

BAU
BAHÇEŞEHİR ÜNİVERSİTESİ

z test

$$t = \frac{\text{difference of sample means}}{\text{standard error of difference of sample means}}$$

$$z = \frac{\text{difference of sample proportions}}{\text{standard error of difference of sample proportions}}$$

The standard error of the difference of two sample proportions is,

$$s_{\hat{p}_1 - \hat{p}_2} = \sqrt{s_{\hat{p}_1}^2 + s_{\hat{p}_2}^2}$$

$$z = \frac{\hat{p}_1 - \hat{p}_2}{s_{\hat{p}_1 - \hat{p}_2}} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{s_{\hat{p}_1}^2 + s_{\hat{p}_2}^2}} \qquad z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{[\hat{p}_1(1 - \hat{p}_1)/n_1] + [\hat{p}_2(1 - \hat{p}_2)/n_2]}}$$

1

2

Pooled variance

- We can pool the observations in the two sample groups to estimate the population standard deviation. $\bar{X}$

- If the hypothesis that the two samples were drawn from the same population is true, then $\hat{p}_1$ and $\hat{p}_2$ are both estimates of the same population proportion p.

Pooled variance

- $\hat{p}_1 = m_1/n_1$, $\hat{p}_2 = m_2/n_2$
- We could consider all the individuals (m_1+m_2) drawn as a sample from a large population of size n_1+n_2 and estimate $\hat{p}$ and s as,

$$\hat{p} = \frac{m_1 + m_2}{n_1 + n_2} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$$

$$s = \sqrt{\hat{p}(1 - \hat{p})}$$

i.e. $n_1\hat{p}_1$ is the number of individuals who died in group 1 if dead/alive is the criteria.

The pooled SE of difference

$$s_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{s^2}{n_1}+\frac{s^2}{n_2}} = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

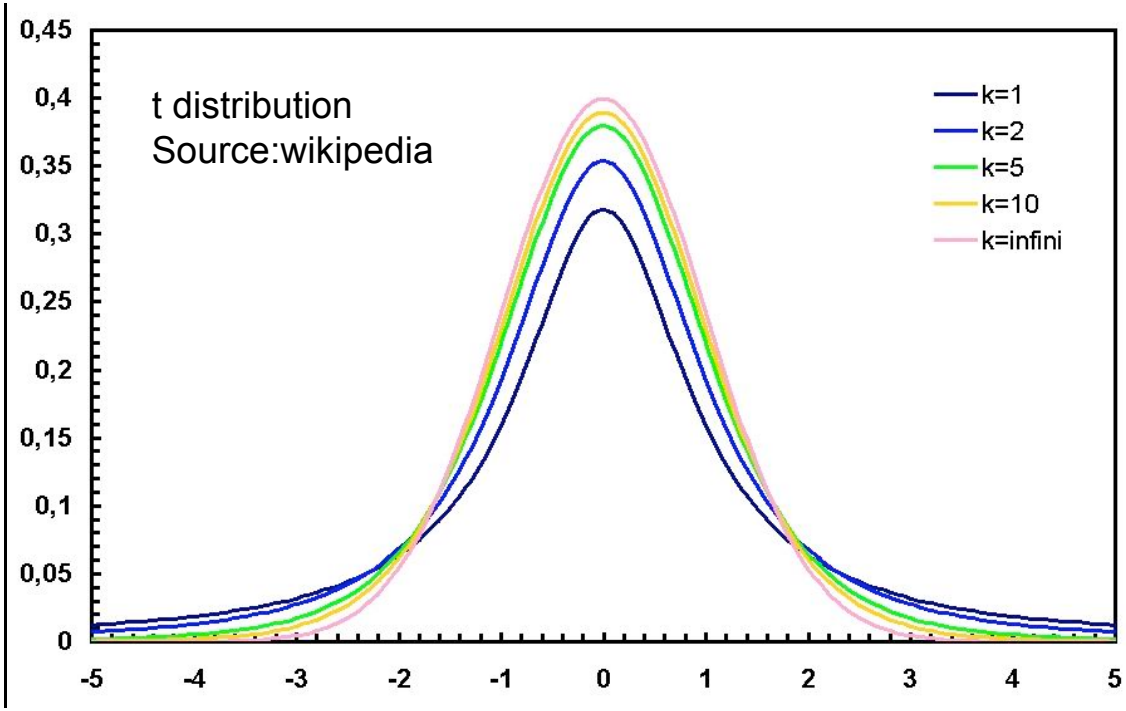
Therefore, z statistic based on a pooled estimate of the standard error of difference in the population proportion is,

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(1/n_1 + 1/n_2)}}$$

Calculate z test statistic, then check Table 4.1 with ∞ degrees of freedom (1.96 for $p<0.05$, 2.58 for $p<0.01$). If z is sufficiently big, we will conclude that the data are inconsistent with the null hypothesis that there is no difference in the proportions.

BAU
BAHÇEŞEHİR ÜNİVERSİTESİ

t distribution



t distribution with infinite degrees of freedom is identical to a normal distribution, so we can use t table with infinite degrees of freedom for the normally distributed z statistic (for large enough
samples).

Yates Correction for Continuity

- z test statistic can only take on discrete values, whereas the theoretical normal distribution is continuous.
- the normal distribution only approximates the actual distribution of the z statistic in a way that yields P values that are always smaller than they should be.
  – conclude treatment had an effect harder.
- Yates correction of continuity modifies z as,

$$z = \frac{|\hat{p}_1 - \hat{p}_2| - {}^1\!/_2(1/n_1 + 1/n_2)}{\sqrt{\hat{p}(1 - \hat{p})(1/n_1 + 1/n_2)}}$$

Example

- Test halothane and morphine are associated with the same mortality rate when used as anesthetic agents in open heart surgery.

- Death rates
  - Halothane – 8/61 (%13.1)
  - Morphine – 10/67 (%14.9)

Example - continued

$$\hat{p} = \frac{8 + 10}{61 + 67} = 0.141$$

np^ for the two samples is; 0.141*61= 8.6, 0.141*67 = 9.4 Both np^ and n(1-p^) exceed 5, so z test can be used.

$$z = \frac{|\hat{p}_{\text{hlo}} - \hat{p}_{\text{mor}}| - \frac{1}{2}(1/n_{\text{hlo}} + 1/n_{\text{mor}})}{\sqrt{\hat{p}(1 - \hat{p})(1/n_{\text{hlo}} + 1/n_{\text{mor}})}}$$

$$= \frac{|0.131 - 0.149| - \frac{1}{2}\left(\frac{1}{61} + \frac{1}{67}\right)}{\sqrt{(0.141)(1 - 0.141)\left(\frac{1}{61} + \frac{1}{67}\right)}} = 0.04$$

Example - continued

Table 4-1
Critical Values of *t* (Two-Tailed) (*continued*)

	Probability of greater value, *P*								
ν	0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
∞	0.6745	1.2816	1.6449	1.9600	2.3263	2.5758	2.8070	3.0902	3.2905

z= 0.04 is too small. p<0.05 cutoff is 1.96.

So, accept the null hypothesis.

There is no difference between halothane and morphine in terms of their mortality rate.

Side note

- According to some results shown in Ch3&4, morphine reduced blood pressure in comparison to halothane. -- process variables are intermediate variables, like change in blood pressure.
- But they did not have any significant differences in terms of their effect to mortality. -- outcome variable.
- Outcome variables are harder to observe, because they require to following the patients for some time and may present difficulty in measurement, like quality of life measures.
- But as much as you can seek evidence that something designed affects patients' OUTCOME.

Another Approach to Testing Nominal Data - Analysis of Contingency Tables

- z-test appropriate for only two attributes or outcomes (ex. dead/alive)
- What if there are more than two possible outcomes/ states?
  - we need sth like ANOVA

Contingency table

Table 5-1 Thrombus Formation in People Receiving Dialysis and Treated with Placebo or Aspirin

	Number of patients		
Sample group	Developed thrombi	Free of thrombi	Treated
Placebo	18	7	25
Aspirin	6	13	19
Total	24	20	44

- 2x2 contingency table
- contingency: possibility, uncertainty
- Table 5.1: most people fell along the diagonal --an association between the absence of aspirin and developing thrombi

No effect scenario

Table 5-2 Expected Thrombus Formation If Aspirin Had No Effect

	Number of patients		
Sample group	Developed thrombi	Free of thrombi	Treated
Aspirin	10.36	8.64	19
Placebo	13.64	11.36	25
Total	24	20	44

- of the 44 people in the study,
  - 25, 25/44 = 57% received placebo
  - 19, 19/44 = 43% received aspirin
  - 24, 24/44= 55% developed thrombi
  - 20, 20/44 = 45% did not
- Table 5.2: experimental results expected if aspirin had no effect when 25 given placebo, 19 given aspirin, 24 of them were destined to develop thrombi regardless of treatment

55% will develop thrombi
if the treatment has no effect

- placebo: 25x55% = 13.64
- aspirin: 19x55% = 10.36

the remaining patients will be free of thrombi

No effect scenario - continued

- Note: Calculate the expected results to two decimal points

- Does 5.1 and 5.2 look similar?
  – No, the actual results look quite different than what we expect of the treatment had no effect.

Another example

Table 5-3 Mortality Associated with Open-Heart Surgery

Anesthesia	Lived	Died	Total no. of cases
Halothane	53	8	61
Morphine	57	10	67
Total	110	18	128

Table 5-4 Expected Mortality with Open-Heart Surgery If Anesthesia Did Not Matter

Anesthesia	Lived	Died	Total no. of cases
Halothane	52.42	8.58	61
Morphine	57.58	9.42	67
Total	110	18	128

Table 5.3 and 5.4 look quite similar. Expected and observed frequencies in each cell are similar. The observation is consistent with no effect of type of anesthesia on mortality.

- 110/128 = 86% lived
- 18/128 = 14% died
- If no effect
  - halothane 61x%86 = 52.42
  - morphine 67x%86 = 57.58

  would live

χ^2 Test Statistic

- *Describes with a single number how much the observed frequencies differ from the frequencies we expect if there is no relationship between the treatments and the outcome*

- *Takes into account if we expect a large number of people to fall in a given cell, a difference of one person between the expected and observed frequencies is less important than in cases where we expect only a few people to fall in the cell*

χ^2 Test Statistic

$$\chi^2 = \text{sum of } \frac{(\text{observed} - \text{expected number of individuals in cell})^2}{\text{expected number of individuals in cell}}$$

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

- sum over all the cells in the contingency table
- if the observed frequencies are similar to the expected frequencies, 2 will be a small number, χ^2 will be a big number otherwise.

BAU
BAHÇEŞEHİR ÜNİVERSİTESİ

χ^2 Test Statistic

- For Tables 5.1 and 5.2,
  χ^2 =7.10
  – The observed and expected values are quite different. Most probably, aspirin had an effect on thrombi formation.

- For Tables 5.3 and 5.4,
  χ^2 =0.09
  – The observed and expected values are quite similar. Most probably, the treatment (halothane or morphine) did not change the mortality rate.

Small Homework

- $\chi^2 = z^2$

  when there are only two samples and two possible outcomes.

Prove it to yourself.

BAU
BAHÇEŞEHİR ÜNİVERSİTESİ

Rules of χ^2

- The expected number of individuals in all the cells must be at least 5.
- The distribution of χ^2 depends on the number of treatments being compared and the number of possible outcomes. The degrees of freedom is,
  - $\nu = (r-1)(c-1)$
    - r: # of rows, c: # of columns
    - for 2x2, $\nu = 1$
- Table 5.7 Critical values for the χ^2 distribution

Yates continuity correction

- χ^2 distribution is continuous, whereas the set of all possible values that χ^2 can take on is not.
- Yates continuity correction is applied as follows when ν=1.

$$\chi^2 = \sum \frac{\left(|O - E| - \frac{1}{2}\right)^2}{E}$$

- Basically, Yates correction reduces 0.5 from each difference in the contingency table, to reduce the χ^2 statistic (or to reduce the significance) and increase the *p* value.
- It makes it harder to detect a significant difference between the observed and expected values.

Chi-Square for More than Two Treatments

Table 5-5 Consult Physician for Menstrual Problem

Group	Yes	No	Total
Controls	14	40	54
Joggers	9	14	23
Runners	46	42	88
Total	69	96	165

Table 5-6 Expected Frequencies of Physician Consultation If Running Did Not Matter

Group	Yes	No	Total
Controls	22.58	31.42	54
Joggers	9.62	13.38	23
Runners	36.80	51.20	88
Total	69	96	165

$$\chi^2 = \frac{(14-22.58)^2}{22.58} + \frac{(40-31.42)^2}{31.42} + \frac{(9-9.62)^2}{9.62} + \frac{(14-13.38)^2}{13.38} + \frac{(46-36.80)^2}{36.80} + \frac{(42-51.20)^2}{51.20} = 9.63$$

69/165 = 42% consulted a physician

Expectation:

– C: 54x42% = 22.58

– J: 23x42% = 9.62

– DR: 88x42% = 36.80

v= (3-1)*(2-1)=2

p<0.01 for χ^2 =9.63

BAU
BAHÇEŞEHİR ÜNİVERSİTESİ

How to Use χ^2 Statistic

- Tabulate data into a contingency table
- Sum the data in columns and rows and calculate the percentages of all individuals falling into each column and row
- Use the percentages to find out the expected number of people in each cell
- Compute χ^2
- Look up Table 5-7 with the corresponding degrees of freedom.
- In 2x2 table, each cell should have at least 5 people. In bigger tables each cell should at least have 1 person, and no more than 20% of them should be less than 5.

Table 5.7

Table 5-7
Critical Values for the χ^2 Distribution

	Probability of greater value P							
ν	.50	.25	.10	.05	.025	.01	.005	.001
1	.455	1.323	2.706	3.841	5.024	6.635	7.879	10.828
2	1.386	2.773	4.605	5.991	7.378	9.210	10.597	13.816
3	2.366	4.108	6.251	7.815	9.348	11.345	12.838	16.266
4	3.357	5.385	7.779	9.488	11.143	13.277	14.860	18.467
5	4.351	6.626	9.236	11.070	12.833	15.086	16.750	20.515
6	5.348	7.841	10.645	12.592	14.449	16.812	18.548	22.458
7	6.346	9.037	12.017	14.067	16.013	18.475	20.278	24.322
8	7.344	10.219	13.362	15.507	17.535	20.090	21.955	26.124

- Based on the degrees of freedom find the cutoff level of χ^2 statistic for various p significance.

BAU
BAHÇEŞEHİR ÜNİVERSİTESİ

Subdividing Contingency Tables

- When χ^2 test statistic is computed for more than two groups
  - It tells whether there is a difference among all the groups like ANOVA.

- We need to isolate which groups differed from which ones using a procedure like the Multiple Comparisons tests.

Joggers vs Runners

Table 5-5 Consult Physician for Menstrual Problem

Group	Yes	No	Total
Controls	14	40	54
Joggers	9	14	23
Runners	46	42	88
Total	69	96	165

Table 5-8 Physician Consultation among Women Joggers and Runners*

Group	Yes	No	Total
Joggers	9 (11.40)	14 (11.60)	23
Runners	46 (43.60)	42 (44.40)	88
Total	55	56	111

*Numbers in parentheses are expected frequencies if the amount of running does not affect physician consultation.

Subdivide the contingency table to include only the joggers and runners, recalculate expected frequencies using only these two groups.

Joggers vs Runners

It is now a 2x2 contingency table. Apply χ^2 statistics with Yates continuity correction.

$$\chi^2 = \sum \frac{\left(|O - E| - \frac{1}{2}\right)^2}{E}$$

$$= \frac{\left(|9 - 11.40| - \frac{1}{2}\right)^2}{11.40} + \frac{\left(|14 - 11.60| - \frac{1}{2}\right)^2}{11.60}$$

$$+ \frac{\left(|46 - 43.60| - \frac{1}{2}\right)^2}{43.60} + \frac{\left(|42 - 44.40| - \frac{1}{2}\right)^2}{44.40} = .79$$

ν=1

p=0.05 cutoff value = 3.841

Joggers and runners are equally likely to visit their physician.

Joggers & Runners vs Controls

Since joggers and runners are so similar to each other, we can combine the two groups into one and compare them against controls.

Table 5-9 Physician Consultation among Women Who Did and Did Not Run*

Group	Yes	No	Total
Controls	14 (22.58)	40 (31.42)	54
Joggers and runners	55 (46.42)	56 (64.58)	111
Total	69	96	165

$\chi^2 = 7.39$ (cutoff = 6.63 for p=0.01) → p<0.01

We have done two tests on the same data, so we need to do a multiple comparison correction. With Bonferoni, compare it against p = 0.05/2 = 0.025, still significant

→ Runners & Joggers are similar but they differed from controls.

BAU
BAHÇEŞEHİR ÜNİVERSİTESİ

Fisher exact test

- χ^2 statistic can be used for 2x2 contingency tables when each cell has a frequency of at least 5.
- In smaller studies (<5 in some cells) Fisher exact test is the appropriate procedure.
- When the sample size is small, it is possible to list all the possible arrangements of the data.
- The total (two tailed) probability of obtaining the observed data or more extreme patterns in the data is the *p* value associated with the data.

BAU
BAHÇEŞEHİR ÜNİVERSİTESİ

Combinatorics

- number of ways k things can be chosen out of a set of n things.

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Example

	men	women	total
dieting	a	b	a+b
not-dieting	c	d	c+d
total	a+c	b+d	n

$$\left|\frac{\binom{dieting}{men}\binom{non-dieting}{men}}{\binom{total}{men}} = \frac{\binom{a+b}{a}\binom{c+d}{c}}{\binom{n}{a+c}} = \frac{\frac{(a+b)!}{a!b!}\frac{(c+d)!}{c!d!}}{\frac{(n)!}{(a+c)!(b+d)!}} = \frac{(a+b)!(c+d)!(b+d)!(a+c)!}{a!b!c!d!n!}\right.$$

Fisher's exact test

$$P = \frac{\frac{R_1!R_2!C_1!C_2!}{N!}}{O_{11}!O_{12}!O_{21}!O_{22}!}$$

Table 5-10 Notation for the Fisher Exact Test

			Row Totals
	O_{11}	O_{12}	R_1
	O_{21}	O_{22}	R_2
Column Totals	C_1	C_2	N

Example

Table 5-11 Reporting of Use of Fisher Exact Test in the *New England Journal of Medicine* and *The Lancet*

	Test Identified?		
Group	Yes	No	Totals
New England Journal of Medicine	1	8	9
The Lancet	10	4	14
Totals	11	12	23

$$P = \frac{\frac{9!14!11!12!}{23!}}{1!8!10!4!} = .00666$$

More extreme pattern

Table 5-12 More Extreme Pattern of Observations in Table 5-11, Using Smallest Observed Frequency (in This Case, 1)

	Test Identified?		
Group	Yes	No	Totals
New England Journal of Medicine	0	9	9
The Lancet	11	3	14
Totals	11	12	23

$$P = \frac{\frac{9!14!11!12!}{23!}}{9!0!3!11!} = .00027$$

Two tailed Fisher exact test

	Yes	No	Total
NEJM	1	8	9
Lancet	10	4	14
Total	11	12	23

	Yes	No	Total
NEJM	0	9	9
Lancet	11	3	14
Total	11	12	23

	Yes	No	Total
NEJM	0	9	9
Lancet	11	3	14
Total	11	12	23

	Yes	No	Total
NEJM	0	9	9
Lancet	12	2	14
Total	12	11	23

More Choices

	Yes	No	Total
NEJM	1	8	9
Lancet	10	4	14
Total	11	12	23

	Yes	No	Total
NEJM	2	7	9
Lancet	9	5	14
Total	11	12	23

	Yes	No	Total
NEJM	3	6	9
Lancet	8	6	14
Total	11	12	23

	Yes	No	Total
NEJM	4	5	9
Lancet	7	7	14
Total	11	12	23

	Yes	No	Total
NEJM	5	4	9
Lancet	6	8	14
Total	11	12	23

	Yes	No	Total
NEJM	6	3	9
Lancet	5	9	14
Total	11	12	23

	Yes	No	Total
NEJM	7	2	9
Lancet	4	10	14
Total	11	12	23

	Yes	No	Total
NEJM	8	1	9
Lancet	3	11	14
Total	11	12	23

	Yes	No	Total
NEJM	9	0	9
Lancet	2	12	14
Total	11	12	23

More Choices

	Yes	No	Total
NEJM	X	X	9
Lancet	1	13	14
Total	11	12	23

	Yes	No	Total
NEJM	X	X	9
Lancet	0	14	14
Total	11	12	23

Two tailed Fisher exact test

- Compute the probability associated with all the possible tables.
- Add all the probabilities that are equal to or smaller than the probability associated with the observed data.
- p=0.00666+0.0027+0.00242+0.00007=0.00944

Association between two variables

Table 5-14 Arrangement of Data to Compute Relative Risk

Sample group	Number of people		
	Disease	No disease	Total
Treated (or exposed to risk factor)	a	b	a + b
Control (or not exposed to risk factor)	c	d	c + d
Total	a + c	b + d	

$$\text{RR} = \frac{\text{Probability of event in } \textit{treatment} \text{ group}}{\text{Probability of event in } \textit{control} \text{ group}}$$

$$\text{RR} = \frac{a/(a+b)}{c/(c+d)}$$

Prospective study versus case-control study

- To compute a relative risk, the data must be collected as a part of a prospective study in which people are randomized to treatment or control or subjects are followed forward in time after they are exposed to the risk factor.
- Case control studies: retrospectively in time
- Unlike prospective studies, case control studies are done after the fact.
  - Cases: People who experienced the outcome
  - Controls: People who did not experience the outcome but are similar to the cases in all relevant ways (age, the exposure to tobacco smoke etc.)

Odds Ratio

$$\text{OR} = \frac{\text{Odds of exposure in } \textit{cases}}{\text{Odds of exposure in } \textit{controls}}$$

Table 5-15 Arrangement of Data to Compute Odds Ratio

Sample group	Number of People: Disease "cases"	Number of People: No disease "controls"
Exposed to risk factor (or treatment)	a	b
Not exposed to risk factor (or treatment)	c	d
Total	a + c	b + d

$$\text{Odds of disease in } \textit{cases} = \frac{a/(a+c)}{c/(a+c)} = \frac{a}{c}$$

$$\text{Odds of disease in } \textit{controls} = \frac{b/(b+d)}{d/(b+d)} = \frac{b}{d}$$

$$\text{OR} = \frac{a/c}{b/d} = \frac{ad}{bc}$$