Statistically Speaking

Seven Fundamental Statistical Tests for Categorical Data

Kim Love

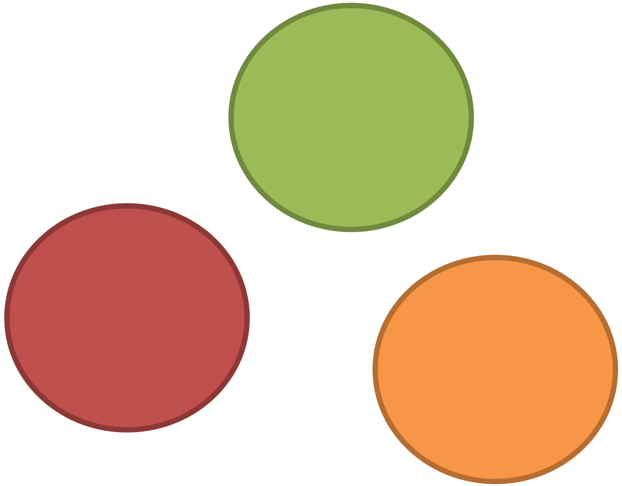
Outline of Topics

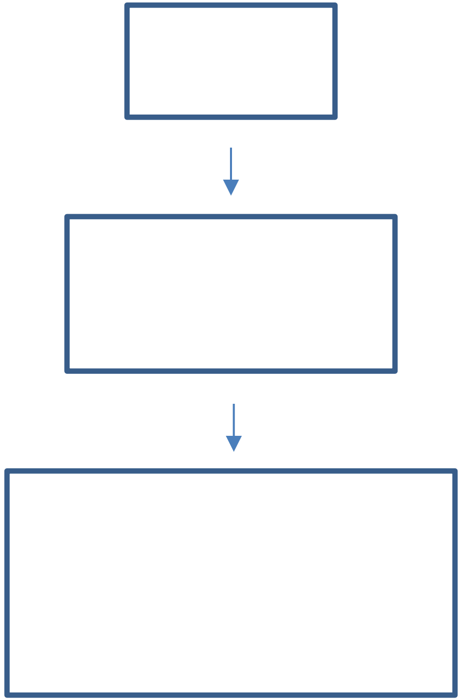
- Introduction
- Goodness of Fit Tests
  - Chi-Square Test
- Tests of Independence
  - Chi-Square Test; Cochran-Mantel-Haenszel Test
  - Note on Measures of Association
- Tests of Homogeneity
  - Chi-Square Test; Fisher Exact Test; Two-Sample Z Test for Proportions; McNemar Test of Symmetry
- Discussion

Introduction

Categorical Data

- Nominal (“color” = red, green, orange)
- Ordinal (“size” = small, medium, large)





Introduction

"Simple" Tests

- Generally for two variables
- Simple questions:
  - Are these measures related?
  - Are these groups different?
- All for nominal data
- Equations ahead!
  - Maybe not be simple on first glance

Introduction

Chi-Square Distributions

- Observed: number of observations *observed* in a category
- Expected: number of observations *expected* in a category

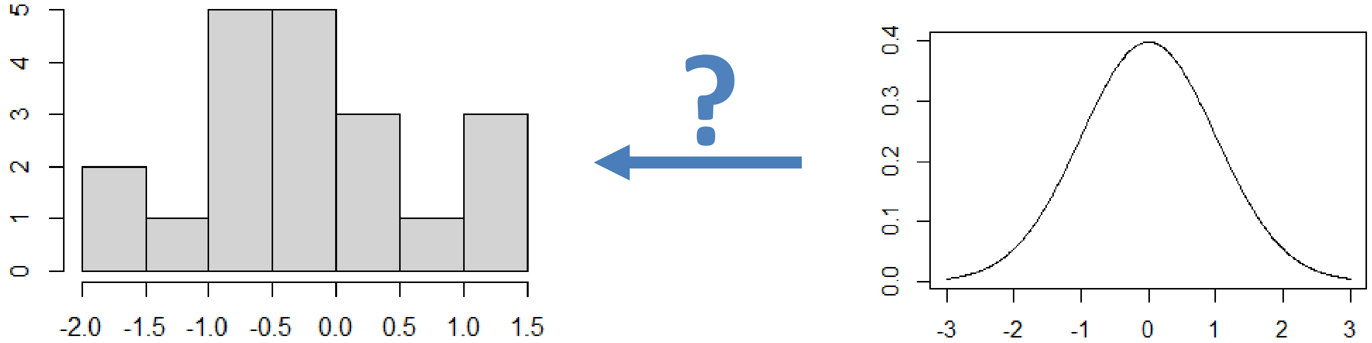
$$\sum \frac{(Observed - Expected)^2}{Expected} = \chi^2$$

- Many variations
- Approximation considered acceptable if
  - Expected in all categories ≥ 1
  - 80% of categories, expected ≥ 5

Goodness of Fit Tests

Goodness of fit

- One variable test
- "Does this distribution fit these data?" or "Could these data be produced by this distribution?"
- Familiar examples: Kolmogorov-Smirnov, Shapiro-Wilk



Goodness of Fit Tests: Chi-Square Test

Example 1: Eye Color (multinomial distribution)

- Eye color alleles: B and b
- Four possible combinations
  - BB = two brown eyes
  - Bb, bB = one brown, one blue eye
  - bb = two blue eyes
- Question: Are both alleles equally likely to be passed to offspring?

Goodness of Fit Tests: Chi-Square Test

Example 1: Eye Color

- 100 (independent) sets of parents, all Bb/bB
- 1 offspring per set of parents
- Eye color recorded:

Two Brown	Mixed	Two Blue	Total
36	43	21	100

- Expected under null hypothesis (equally likely B/b combinations)?
  - 25% each: BB, Bb, bB, bb

Two Brown	Mixed	Two Blue	Total
25	50	25	100

Goodness of Fit Tests: Chi-Square Test

Example 1: Eye Color

- Chi-square statistic:

$$\sum_{i=1}^{k} \frac{(Observed_i - Expected_i)^2}{Expected_i} = \chi^2_{k-1}$$

- k = 3 categories

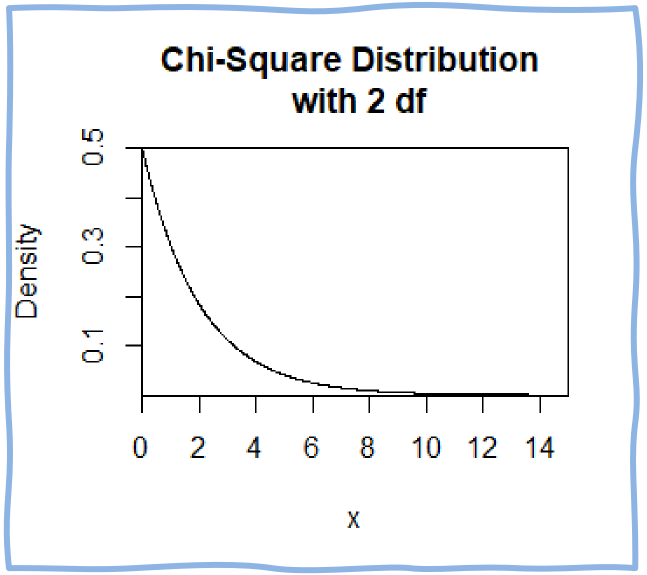
Goodness of Fit Tests: Chi-Square Test

Example 1: Eye Color

- Chi-square statistic:

$$\sum_{i=1}^{3} \frac{(Observed_i - Expected_i)^2}{Expected_i} = \chi_2^2$$

- Calculation: $\frac{(36-25)^2}{25} + \frac{(43-50)^2}{50} + \frac{(21-25)^2}{25}$ = 6.46
- Determine: p value = P(χ_2^2) > 6.46 = 0.0396



Goodness of Fit Tests: Chi-Square Test

Example 2*: Number of Shoppers per Minute (Poisson Distribution)

- Number of shoppers entering store in one minute
- Question: Is this Poisson distributed (potentially with a mean of 2 per minute)?

*Borrowed from Prof. Robert Schulman's Inference Fundamentals course

Goodness of Fit Tests: Chi-Square Test

Example 2: Number of Shoppers per Minute

- 200 randomly selected minutes in store observed
- Number of people entering during each minute recorded

	Number of People						
	0	1	2	3	4	5	≥ 6
Number of Minutes	18	44	49	43	27	12	7
Expected under Poisson (λ=2)	27.07	54.13	54.13	36.09	18.05	7.22	3.31

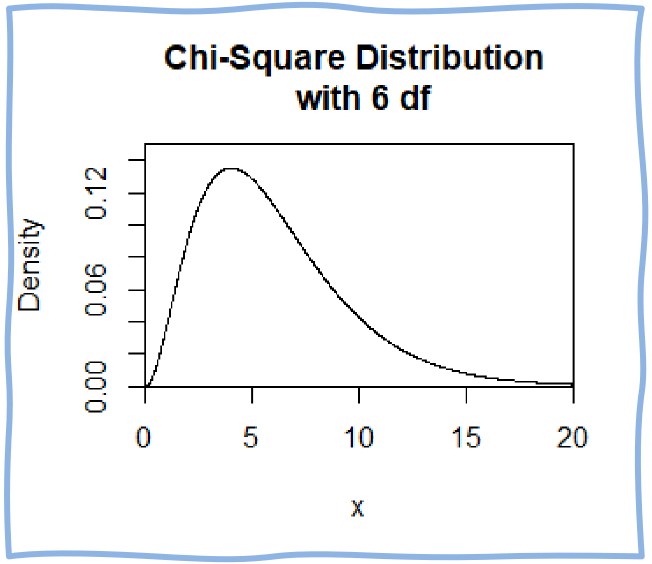
Goodness of Fit Tests: Chi-Square Test

Example 2: Number of Shoppers per Minute

$$\sum_{i=1}^{7} \frac{(Observed_i - Expected_i)^2}{Expected_i} = \chi_6^2$$

- Calculation: $\frac{(18-27.07)^2}{27.07} + \frac{(44-54.13)^2}{54.13} + \cdots + \frac{(7-3.31)^2}{3.31}$ = 18.46

- P(χ_6^2) > 18.46 = 0.0024



Goodness of Fit Tests: Chi-Square Test

Example 2: Number of Shoppers per Minute

- Rejected; is Poisson rejected? Or just λ=2?
- Estimate λ from data: try λ = 2.41

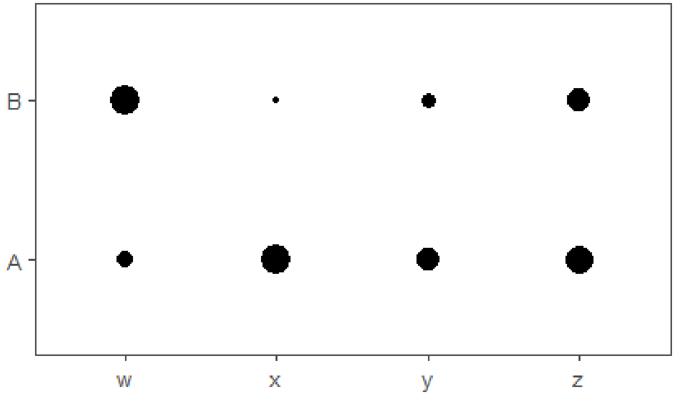
	Number of Minutes						
	0	1	2	3	4	5	≥ 6
Number of People	18	44	49	43	27	12	7
Expected under Poisson (λ=2.41)	17.96	43.29	52.17	41.91	25.25	12.17	7.26

- Calculation: $\frac{(18-17.96)^2}{17.96} + \frac{(44-43.29)^2}{43.29} + \cdots + \frac{(7-7.26)^2}{7.26}$ = 0.365
- P(χ^2_5) > 0.365 = 0.9962

Tests of Independence

Independence

- Two variable test—mostly
- “Are the values of these two variables related?” or “Does the value of one variable help predict the value of the other?”
- Not the same concept as correlation, but connected



Tests of Independence: Chi-Square Test

- Recall recurring chi-square theme:

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

- What does this look like for independence?

Tests of Independence: Chi-Square Test

Example: Car Transmission and Drive Train Types

- Is manual transmission availability related to a car's drive train type?

Manual Transmission Available?	Drive Train: Front	Rear	4WD	Total
No	22	7	3	32
Yes	45	9	7	61
Total	67	16	10	93

Tests of Independence: Chi-Square Test

Example: Car Transmission and Drive Train Types

Manual Transmission Available?	Drive Train: Front	Rear	4WD	Total
No	22	7	3	32
Yes	45	9	7	61
Total	67	16	10	93

- Under independence, how many expected manual transmission x front wheel drive models?
  - P(MT) = 61/93 = 65.59%; P(Front) = 67/93 = 72.04%
  - P(MT & Front) = 65.59% x 72.04% = 47.25%
  - Expected count of MT & Front = 47.25% x 93 = 43.94

Tests of Independence: Chi-Square Test

Example: Car Transmission and Drive Train Types

- Chi-square statistic:

$$\sum_{j=1}^{3}\sum_{i=1}^{2}\frac{\left(Observed_{ij}-Expected_{ij}\right)^2}{Expected_{ij}}=\chi^2_{(2-1)(3-1)}$$

- Calculation: $\frac{(22-23.05)^2}{23.05}+\frac{(45-43.95)^2}{43.95}+\cdots+\frac{(7-6.56)^2}{6.56}$ = 0.778
- P(χ^2_2) > 0.778 = 0.678

Tests of Independence: Cochran-Mantel-Haenszel Test

- Similar to chi-square test
  - Includes a “blocking” or “strata” variable
  - Are two variables independent after stratification?

Tests of Independence: Cochran-Mantel-Haenszel Test

Example: Car Transmission and Drive Train Types—by Origin

- Previous result indicates independence
- Does independence hold if we break into strata?

Foreign Cars

Manual Transmission Available?	Drive Train: Rear/4WD	Drive Train: Front	Total
No	3	3	6
Yes	9	30	39
Total	12	33	45

US Cars

Manual Transmission Available?	Drive Train: Rear/4WD	Drive Train: Front	Total
No	7	19	26
Yes	7	15	22
Total	14	34	48

Tests of Independence: Cochran-Mantel-Haenszel Test

Example: Car Transmission and Drive Train Types—by Origin

- A single stratum 2 x 2 table:

Manual Transmission Available?	Drive Train	
	Rear/4WD	Front
No	a	b
Yes	c	d

- $n = a + b + c + d$ (total observations within a stratum)

Tests of Independence: Cochran-Mantel-Haenszel Test

Example: Car Transmission and Drive Train Types—by Origin

- Chi-square statistic:

$$\frac{\left\{\left|\sum\left(a-\frac{(a+b)(a+c)}{n}\right)\right|-0.5\right\}^2}{\sum\left(\frac{(a+b)(a+c)(b+d)(c+d)}{n^3-n^2}\right)}=\chi_1^2,$$

where sums are over all strata

Tests of Independence: Cochran-Mantel-Haenszel Test

Example: Car Transmission and Drive Train Types—by Origin

Foreign Cars

Manual Transmission Available?	Drive Train: Rear/4WD	Drive Train: Front	Total
No	3	3	6
Yes	9	30	39
Total	12	33	45

US Cars

Manual Transmission Available?	Drive Train: Rear/4WD	Drive Train: Front	Total
No	7	19	26
Yes	7	15	22
Total	14	34	48

- Calculation: $\dfrac{\left\{\left(3-\frac{(6)(12)}{45}\right)+\left(7-\frac{(26)(14)}{48}\right)-0.5\right\}^2}{\frac{(6)(12)(39)(33)}{45^3-45^2}+\frac{(26)(14)(22)(34)}{48^3-48^2}} = 0.028$

- $\mathrm{P}(\chi_1^2) > 0.028 = 0.837$

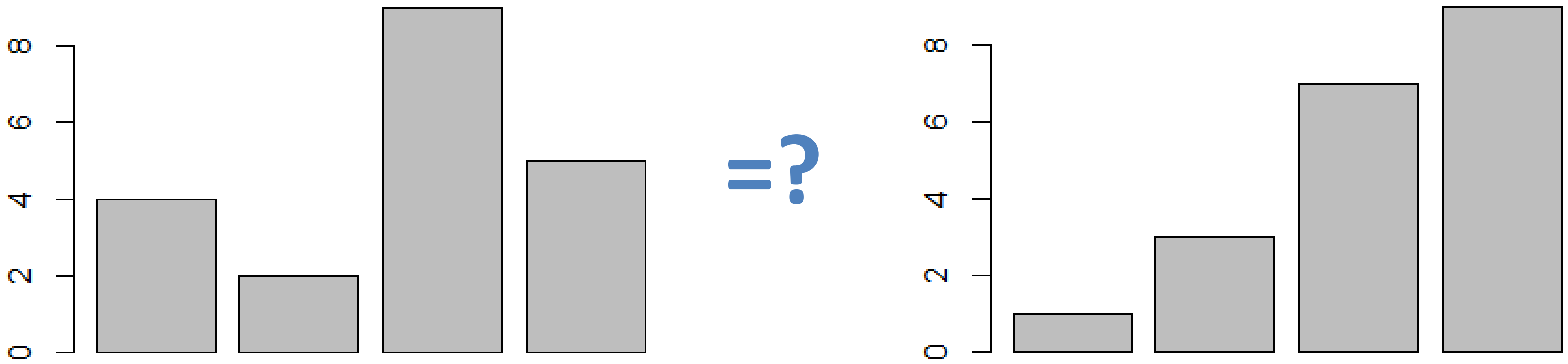
Tests of Independence: Note on Measures of Association

- Tests provide evidence of non-zero effect (when it exists)
  - *Strength* of effect also important
  - *Direction* of effect important for ordinal variables
- See https://programs.theanalysisfactor.com/statistically-speaking/trainings/measures-of-association-beyond-pearsons-correlation/

Tests of Homogeneity

Homogeneity

- Two variable test
- "Are these populations the same with respect to this variable?" or "Is the distribution of this variable the same within these populations?"
- Like t-tests, but for categories



Tests of Homogeneity: Chi-Square Test

Example: Reasons for Immigration in Madagascar

- Is the distribution of reasons for immigrating to an urban center the same for male and female residents of Madagascar?

	Reason for Move				
Gender	Education	Family	Work/Money	Other	Total
Female	4	11	10	7	32
Male	11	2	7	3	23
Total	15	13	17	10	55

Tests of Homogeneity: Chi-Square Test

Example: Reasons for Immigration in Madagascar

- Recurring chi-square statistic theme:

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

- If male/female are the same, for example:
  - Percent of people naming education = 15/55 = 27.27%
  - Expected education count for female: 27.27% x 32 = 8.73
  - Expected education count for male 27.27% x 23 = 6.27

Tests of Homogeneity: Chi-Square Test

Example: Reasons for Immigration in Madagascar

- In fact, chi-square statistic turns out to be

$$\sum_{j=1}^{2}\sum_{i=1}^{4}\frac{\left(Observed_{ij}-Expected_{ij}\right)^2}{Expected_{ij}}=\chi^2_{(4-1)(2-1)},$$

same as if testing for independence

- Calculation: $\frac{(4-8.73)^2}{8.73}+\frac{(11-6.27)^2}{6.27}+\cdots+\frac{(3-4.18)^2}{4.18}$ = 10.433
- P(χ^2_3) > 10.433 = 0.015

Tests of Homogeneity: Fisher Exact Test

- Earlier: chi-square approximation considered acceptable if
  - Expected always ≥ 1
  - 80% of expected ≥ 5
- Fisher exact test does not rely on any approximation
  - Appropriate for small sample sizes
  - Appropriate for “rare event” data

Tests of Homogeneity: Fisher Exact Test

Example: Reasons for Immigration in Madagascar

- Suppose we limit reasons to family/not family and ask again: Is the distribution of reasons for immigrating to an urban center the same for male and female residents of Madagascar?

	Reason for Move		
Gender	Family	Other reason	Total
Female	11	21	32
Male	2	21	23
Total	13	42	55

Tests of Homogeneity: Fisher Exact Test

Example: Reasons for Immigration in Madagascar

- Fix margins of table:

	Reason for Move		
Gender	Family	Other reason	Total
Female	11	21	32
Male	2	21	23
Total	13	42	55

- Assume male and female equally likely to move for family
- Given 13/55 moved for family, and 32/55 are female, calculate

$$P(11 \text{ female}, 2 \text{ male}) = 0.0225$$

Tests of Homogeneity: Fisher Exact Test

Example: Reasons for Immigration in Madagascar

- What about even more extreme situations, if male and female equally likely to move for family?

P(12 female, 1 male) = 0.0036, P(13 female, 0 male) = 0.0002, …

Tests of Homogeneity: Fisher Exact Test

Example: Reasons for Immigration in Madagascar

- What about even more extreme situations, if male and female equally likely to move for family?

P(12 female, 1 male) = 0.0036, P(13 female, 0 male) = 0.0002,

P(0 female, 13 male) ≈ 0, …, P(4 female, 9 male) = 0.0202

- Final p value = 0.0225 + 0.0036 + 0.0002 + 0 + … + 0.0202 = 0.0510

Tests of Homogeneity: Two-Sample Z Test for Proportions

- For homogeneity, chi-square is approximate
- We can also approximate a Z statistic (i.e., standard normal) when table 2 x 2
- Ultimately, they are equivalent
  - $(Z)^2 = \chi_1^2$
  - p value from Z test = p value from chi-square test

Tests of Homogeneity: Two-Sample Z Test for Proportions

Example: Reasons for Immigration in Madagascar

	Reason for Move		
Gender	Family	Other reason	Total
Female	11	21	32
Male	2	21	23
Total	13	42	55

- If P(reason = family) the same, best estimate = 13/55 = 23.63%
- P(reason = family) for females = 11/32 = 34.37%
- P(reason = family) for males = 2/23 = 8.69%

Tests of Homogeneity: Two-Sample Z Test for Proportions

Example: Reasons for Immigration in Madagascar

- Calculation: $Z = \frac{0.3437-0.0869}{\sqrt{0.2363(1-0.2363)\left(\frac{1}{32}+\frac{1}{23}\right)}} = 2.21$
- $P(Z > |2.21|) = 0.027$
- Note: Z can be "one tailed," chi-square cannot
  - Are women more likely to move for family?
  - $P(Z > 2.21) = 0.014$

Tests of Homogeneity: McNemar Test of Symmetry

- What if values of variables not independent?
  - Example: pre/post
- This requires a different test
- Similar to a paired t-test

Tests of Homogeneity: McNemar Test of Symmetry

Example: Exercise Program and Pre-Diabetic Status

- Is the probability of pre-diabetes the same before and after an exercise program?

	Post-Exercise		
Pre-Exercise	Not Pre-Diabetic	Pre-Diabetic	Total
Not Pre-Diabetic	15	2	17
Pre-Diabetic	10	8	18
Total	25	10	35

Tests of Homogeneity: McNemar Test of Symmetry

Example: Exercise Program and Pre-Diabetic Status

- Table tells us:
  - Pre-exercise, 18/35 = 51.42% are pre-diabetic
  - Post-exercise, 10/35 = 28.57% are pre-diabetic
  - Overall, 15 + 8 = 23 did not change status; 10 improved, 2 declined
- McNemar test: for those who change, assume P(improve) = P(decline) = 50%

Tests of Homogeneity: McNemar Test of Symmetry

Example: Exercise Program and Pre-Diabetic Status

- For small samples, do an exact test (very similar to Fisher's)
- For larger samples, use standard normal approximation:

$$\frac{n_+ - n_-}{\sqrt{n_+ + n_-}} = Z$$

- Calculation: $\frac{10-2}{\sqrt{10+2}}$ = 2.309
- P(Z > |2.309|) = 0.0209

Discussion

- Questions?

Resources at The Analysis Factor

Statistically Speaking Presentations: Log in at

https://programs.theanalysisfactor.com

Go to your Statistically Speaking Membership, click on Trainings, and search for:

- Measures of Association: Beyond Pearson's Correlation
- Non-Parametric Analyses
- Determining Levels of Measurement: What Lies Beneath the Surface
- Analysis of Ordinal Variables — Options Beyond Nonparametrics
- Types of Regression Models and When to Use Them

References

- Practical Nonparametric Statistics by W. J. Conover: click