Lecture 15: Chi-Square Tests

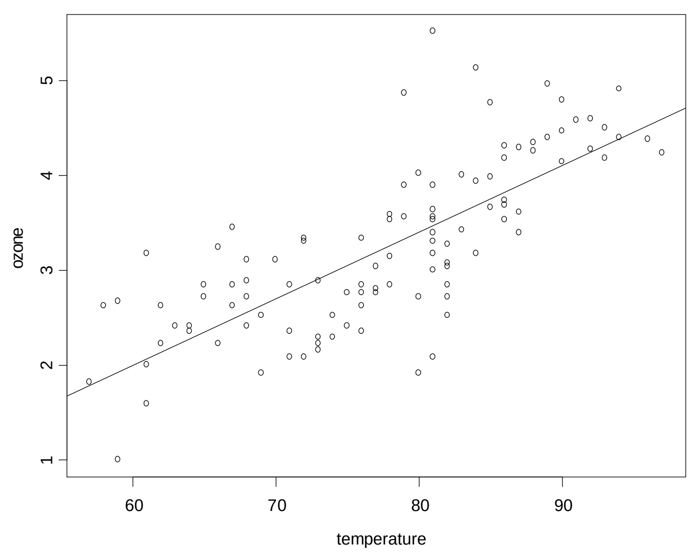
• 15. 1 Parametric and Non-Parametric Methods

- Parametric regression methods are dictated by a distributional form of model errors, which allow a strict mathematical distribution
    - Parametric data have an underlying normal distribution, which allows for more conclusions to be drawn as the shape can be mathematically described. Non-normal distributions require non-parametric methods

- One basic distinction for parametric versus non-parametric is
    - If the measurement scale is interval or ratio → parametric statistics
    - If the measurement scale is nominal or ordinal → non-parametric statistics

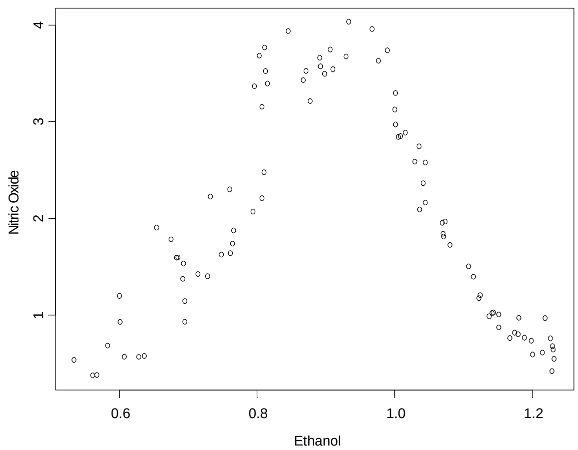
- Non-parametric statistical procedures are less powerful because they use less information in their calculation
    - For example, a parametric correlation (Pearson) uses information about the mean and deviation from the mean while a non-parametric correlation (Spearman) will use only the ordinal position of pairs of scores
    - Computation of *t* involves using sample substitutes for the population standard error and mean

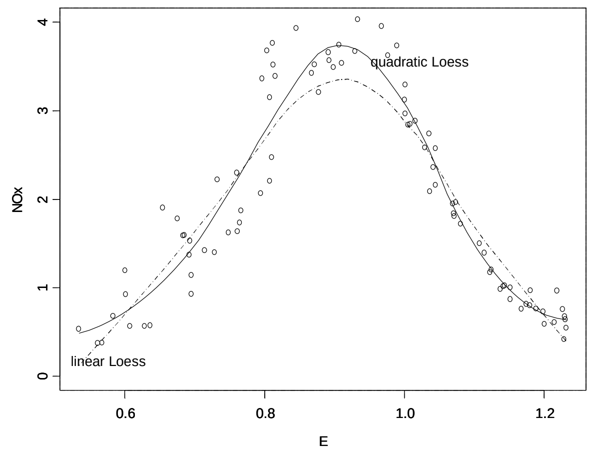
- Correlation, regression, ANOVA, and *t*-tests are all *parametric methods* because they all involve estimation of a parameter with a statistic and rely on distributional assumptions

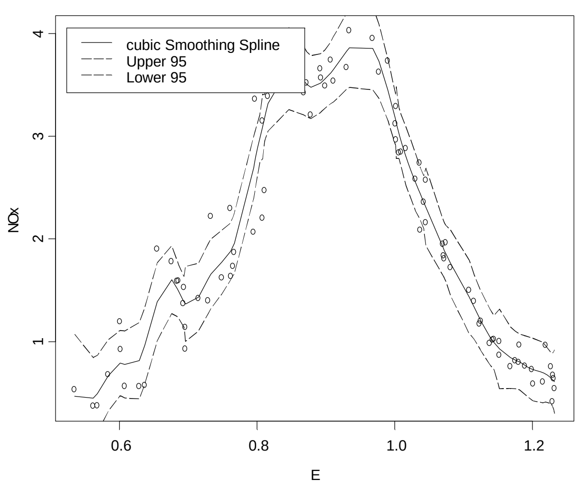


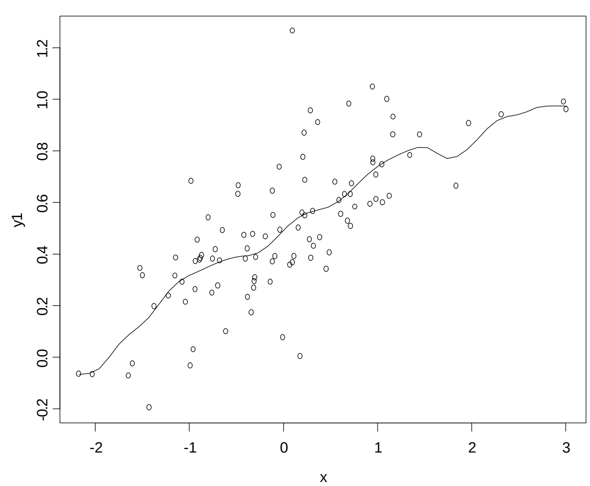
- Robust Regression models relax the strict assumption, thus, allowing fuzzy distributional spaces that are more realistic yet still approximate the parametric model

- Non-Parametric methods ignore the distribution form and allow all dimensions of all possible residual distributions
    - Some methods do not depend on estimation of population values
    - Useful with count data, or data that are nominal or ordinal rather than interval or ratio



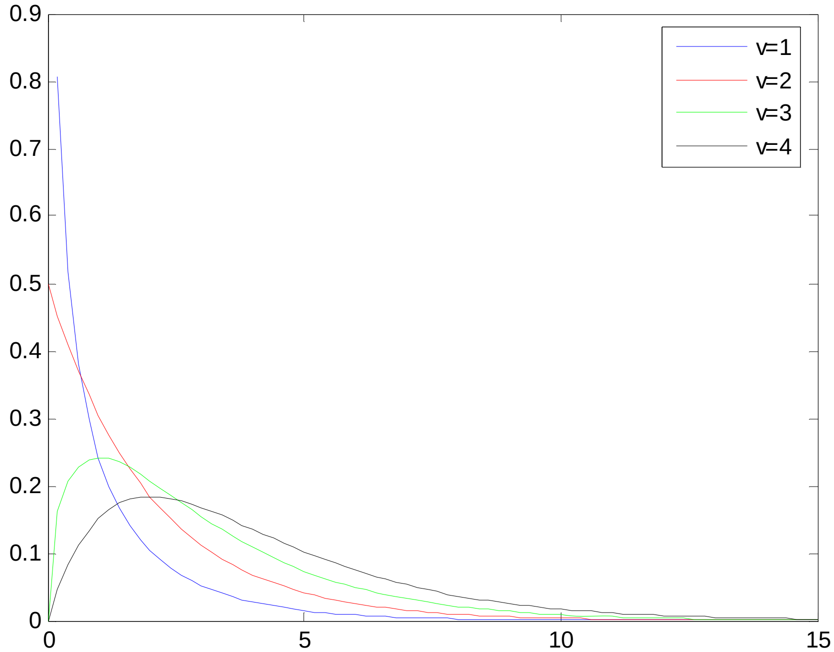






- Chi-square tests fall somewhere in between: useful with count data, but still involve reference to a hypothetical distribution

- **15. 2 The Chi-Square Distribution**



- It is a function with only one parameter (υ), the degrees of freedom

  - In all cases, mean = υ and variance = 2 υ

- It is "chi-square" because the χ^2 distribution is a distribution of squared *z*-scores

  - Let be a normal distribution $N(\mu, \sigma^2)$ from which to sample one observation x_1 and calculate z_{x1}^2

$$z_{x1}^2 = \frac{(x_1 - \mu)^2}{\sigma^2}$$

- If we repeat this procedure an infinite number of times and plot the distribution of z_{xi}^2, the resulting distribution will look exactly like the χ^2 distribution on 1 *df*. In fact,

$$\chi_{(1)}^2 = z^2$$

  - It is a positively skewed distribution
  - Minimum value is zero
  - Maximum value is ∞
  - Most values are between zero and 1; most around zero

- What if we took 2 values of z^2 at random and added them?

$$z_1^2 = \frac{(y_1 - \mu)^2}{\sigma^2}; \; z_2^2 = \frac{(y_2 - \mu)^2}{\sigma^2}$$

$$\chi^2_{(2)} = \frac{(y_1 - \mu)^2}{\sigma^2} + \frac{(y_2 - \mu)^2}{\sigma^2} = z_1^2 + z_2^2$$

  - Same minimum and maximum as before, but now average should be a bit bigger

  - Chi-square is the distribution of a sum of squares. Each squared deviation is taken from the unit normal: $N(0,1)$. The shape of the chi-square distribution depends on the number of squared deviates that are added together

- The expected value chi-square is *df*

- The mean of the chi-square distribution is its degrees of freedom

- The expected variance of the distribution is 2*df*

- If the variance is 2*df*, the standard deviation must be $\sqrt{2df}$

- Chi-square is additive

$$\chi^2_{(v_1+v_2)} = \chi^2_{(v_1)} + \chi^2_{(v_2)}$$

- For any sampled *N* scores at a time, the resulting distribution of $z^2{}_N$ is a chi-square distribution of the form

$$\chi^2{}_N = \sum_{i=1}^{N} z_i{}^2 = \sum \frac{(X_i - \mu)^2}{\sigma^2}$$

- ## 15. 3 Chi-Square Test: Example – Frequencies

- Suppose we want to examine the fairness of the admittance rate for males and females at UCD

- Because we don't have access to admittance records, we sit on a bench and record the sex of everyone who walks by in an hour

Male	*Female*
146	152

- These frequencies are not equal, but we wouldn't expect them to be. The question is, do they depart significantly from equality by more than chance levels?

- If UCD had a "fair" admittance policy, we would expect 149 men and 149 women

	Male	*Female*	*Total*
Observed	146	152	298
Expected	149	149	298

- Is the UCD population split 50/50 for men/women?

- H_0: *The relative frequencies of occurrence of observed events follow a specified frequency distribution* (in this case, uniform)

- Calculate χ^2

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

- Chi-square is the sum of the squared differences between observed and expected frequencies

  - It would equal zero if *O* and *E* were exactly the same within each cell

- In this example

$$\chi^2 = \sum \frac{(O-E)^2}{E} = \underbrace{\frac{(146-149)^2}{149}}_{male} + \underbrace{\frac{(152-149)^2}{149}}_{female} = 2(.0604) = .1208$$

- For the chi-square test of goodness of fit, $df = C - 1$, where C = number of categories

    - Thus, $df = 2 - 1 = 1$

    - Suppose we choose $\alpha = .05$

    - Critical $\chi^2 = 3.84$

- Observed χ^2 < Critical χ^2. Hence, we *fail to reject* H_0. We conclude that 146 and 152 are not far enough away from 149 to claim that UCD's admittance policy is unfair

- **15. 4 Example b (> 2 Cells)**

- The chi-square test is not limited to two cells

- Suppose you are interested in studying children's playing preferences in playgrounds

- You observe 51 children in playgrounds and annotate each child's playing preferences (i.e., alone, with others, or both)

	Alone	*Others*	*Both*	*Total*
Observed	7	26	18	51
Expected	17	17	17	51

- If there is no preference, then we might expect to see equal time in all 4 categories (i.e., 17)

$$\chi^2 = \sum \frac{(O-E)^2}{E} = \underbrace{\frac{(7-17)^2}{17}}_{alone} + \underbrace{\frac{(18-17)^2}{17}}_{equal} + \underbrace{\frac{(26-17)^2}{17}}_{others} = \frac{100+1+81}{17} = 10.7$$

- $df = 2$

- $\alpha = .05$

- Critical $\chi^2 = 5.99$, so we *reject* H_0. The observed frequencies *do* differ significantly from equality

• 15. 5 Example c (> Non-Uniform Expected Frequencies)

- The chi-square test is not limited to uniform expected frequencies

- Suppose you are interested in studying whether there are racial disparities in the populations of California prisons

- A selection of random 300 prison records yields the following data

African-American	*Caucasian*	*Hispanic*	*Other*	*Total*
70	50	168	12	300

- Based on census data, we know that the population in California has the following racial-ethnic makeup: 20% African-American, 20% non-Hispanic Caucasian, 57% Hispanic, and 3% other

- This implies that we should expect our sample to have the following distribution: 60 African-Americans, 60 Caucasian, 171 Hispanics, and 9 Others

	African-American	*Caucasian*	*Hispanic*	*Other*	*Total*
Observed	70	50	168	12	300
Expected	60	60	171	9	300

- But, there seems to be a disparity between our observations and the expectations

$$\chi^2 = \sum \frac{(O-E)^2}{E} = \underbrace{\frac{(70-60)^2}{60}}_{Af-Am} + \underbrace{\frac{(50-60)^2}{60}}_{Cauc} + \underbrace{\frac{(168-171)^2}{171}}_{Hisp} + \underbrace{\frac{(12-9)^2}{9}}_{Other} = \frac{100}{60} + \frac{100}{60} + \frac{9}{171} + \frac{9}{9} = 4.4$$

- $df = 3$

- $\alpha = .05$

- Critical $\chi^2 = 7.82$, so we *fail to reject* H_0. We do not have enough evidence to conclude that there are racial disparities in the prisons population of California

• 15. 6 Chi-Square Test of Independence

- Suppose we want to examine whether two variables (e.g., *ethnicity* and *playing preferences* in the previous examples) are independent of one another

- This situation calls for a chi-square *test of independence*

- Hypotheses:

H_0: X is independent of Y
H_1: X is *not* independent of Y

- Example: Suppose you want to know if people's position on the abortion issue is independent of political affiliation

- You collect data from a random sample of 170 individuals and place them into 4 categories depending on their responses:

	Republican	*Democrat*	*Total*
Pro-Life	73	29	102
Pro-Choice	17	51	68
	90	80	170

- This is a 2 x 2 *contingency table*. We can provide marginal information for rows and columns and obtain *observed* frequencies

- To conduct a chi-square test, we need to calculate the *expected* frequencies (Row total x Column total / N)

	Republican	*Democrat*	*Total*
Pro-Life	54	48	102
Pro-Choice	36	32	68
	90	80	170

- And we can now perform a chi-square test of independence, using expected frequencies (Row total x Column total / *N*)

$$E_{ij} = \frac{R_i C_j}{N}$$

$$E_{11} = \frac{(90)(102)}{170} = 54$$

- The chi-square test of independence is calculated exactly like the test of frequencies (or goodness of fit):

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

$$\chi^2 = \sum \frac{(O-E)^2}{E} = \underbrace{\frac{(73-54)^2}{54}}_{PL,Rep} + \underbrace{\frac{(29-48)^2}{48}}_{PL,Dem} + \underbrace{\frac{(17-36)^2}{36}}_{PC,Rep} + \underbrace{\frac{(51-32)^2}{32}}_{PC,Dem} = \frac{361}{54} + \frac{361}{48} + \frac{361}{36} + \frac{361}{32} = 35.5$$

$$df = (R-1)(C-1) = (2-1)(2-1) = 1$$

- At $df = 1$ and $\alpha = .05$, the critical value is 3.84. Reject H_0 and conclude that abortion stance is *not* independent of political affiliation (i.e., they are related)

- The chi-square test is not limited to 2-way tables, it can be applied to multiway tables

$$E_{ijk} = \frac{R_i C_j T_k}{N}$$

$$df = (R-1)(C-1)(T-1)$$

- ## 15. 7 Advantages/Requirements and Limitations

- Requirements
  - Random sampling
  - Independent observations
  - Mutually exclusive measurement classes
  - Each and every sample observation falls into one category or class interval only
  - Sample *N* is large; average cell frequency should be ≥ 5
  - Can use any level of data
  - It does not involve assumptions about the shape of the population distribution from which the sample is drawn

- Limitations
  - The chi-square test is not appropriate if any of the *expected* cell counts is < 5
    - But this is conservative. Some statisticians claim that chi-square is always inappropriate if any expected frequencies is below 1 or if more than 20% of the cells have cell frequencies below 5
  - The chi-square test is inappropriate if the observations are not independent
    - For example, data from husbands and wives, or children and parents, or repeated measures into the design