Lecture 22

Fisher Exact Test, Non-Parametric Tests for Contingency Tables and Hypothesis Test summary
The Fisher Exact test is a method for dealing with contingency table data that is most commonly done for small tables, 2×2 specifically, but can be done for any size table. It is an extension of the multivariate hypergeometric distribution. The test is to determine if there are non-random associations between the two variables.

The total observations N is equal to the sum of the row sums R_i (for the sum of the ith row), and the sum of the column sums C_j (for the sum of the jth column. The conditional probability of obtaining this matrix given these sums is given by

$$P_{cutoff} = \frac{\prod(R_i!)\prod(C_j!)}{N!\prod(a_{ij}!)}$$

For example, for a 2×2 table,

	Math. Mag.	Science	
math	5	0	$R_1 = 5$
biology	1	4	$R_2 = 5$
	$C_1 = 6$	$C_2 = 4$	$N = 10.$

Computing P_{cutoff} gives

$$P_{\text{cutoff}} = \frac{5!^2\, 6!\, 4!}{10!\,(5!\,0!\,1!\,4!)} = 0.0238,$$

Now, consider the other values a table of this size could take with the same row and column sums. There are only 4 total.

$$\begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix} P = 0.2381 \qquad \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} P = 0.2381$$

$$\begin{bmatrix} 3 & 2 \\ 3 & 2 \end{bmatrix} P = 0.4762 \qquad \begin{bmatrix} 1 & 4 \\ 5 & 0 \end{bmatrix} P = 0.0238,$$

We’ve calculated the associated probabilities using the same formula. These probabilities add to 1. Since the P-value of the table we have is less than 0.05, it is statistically significant.

Fortunately, the fisher.test() function can conduct this test for us. We provide the test data in a table format just as we needed to for the χ^2 test of independence.

For the Fisher Exact test, the total sample sizes should be small and more than 20% of cells have a count less than 5. For larger sample sizes, we are essentially using a continuous approximation and we don't need this test. As the samples get large, the factorials can also tend to blow up.

We can also do non-parametric tests on contingency tables, in particular, randomization tests. The general procedure is to fix the sum of the marginal probabilities, and then sample to fill the table for N total observations. Calculate the χ^2 value for the resulting table as your test statistic. If the sample size is very large, it will become difficult to list all the possible tables (even if, as with the Fisher Exact test, we fix the row and column sums). So, sampling thousands of such tables will produce a distribution (that approximates a χ^2 distribution), from which we can compute the P-value compared to our original data. While we can certainly write the code for this test ourselves, there is an R package that can perform this test for us, and we'll look at it in the final lab.

We can further extend our contingency table if we wanted to three-way tables, but this is less common.

Review of Hypothesis Testing – which test do we use?
As we saw when we looked at statistical graphs, hypothesis testing depends on many factors to choose the correct test. Different graphs display information, sometimes the same information, in different ways. Some graphs are designed for very specific types of data. Some graphs convey more or less information about the data. Specific graph types may be used for very specific scenarios. Likewise, for hypothesis tests: tests have very specific types of data to work on, in different formats, and using different assumptions.

You could create a decision chart for deciding which test to use.

Is the variable categorical (the measure is counts or proportions)? Or is it numerical?

How many variables are involved? Is it one variable or two or more? What parameter (if any) are you testing?

	One Variable	Two Variables	Three or More Variables
Categorical (or maybe Discrete)	Proportion Test Goodness-of-Fit Test	Two-Sample Proportion Test χ^2 test Fisher Exact Test	χ^2 test Fisher Exact Test
Both	N/A	Two-sample T-Test (Pooled) Two-sample T-Test (unpooled) Wilcoxon rank-sum test ANOVA (one way)	ANOVA (N-way) Friedman's ANOVA
Numerical	T-Test Z-Test Poisson (χ^2) Variance (F) Wilcoxon test Goodness-of-Fit Test	Regression (next semester) Paired T-test Wilcoxon test	Regression (next semester) ANOVA

Permutation tests can be designed for most of these situations, but it's a good idea to figure out the best parametric test first, and then consider the corresponding non-parametric test. Unless it's something like a median, which does not have a standard parametric test.
Check your assumptions. What distribution is appropriate for the test, if any? (Do you have what you need? You may need to test this first.) Do you have a small sample or a large one? Calculating P-values or the test procedure itself may depend on the sample size.

Sometimes you can tell what kind of test you need to do (or what tests you can eliminate) based on the way the data is provided to you. Paired t-tests can only be conducted from raw data, for instance, with equal size samples, while other kinds of t-tests can be done with sample summary statistics.

If you have difficulty selecting the correct test, it may help you to make your own list of things to consider and look for, maybe even a decision tree. Thinking through these things carefully in advance can help you ensure that the test you conduct is able to give you meaningful information.

Next time, we'll say a little bit about a Bayesian approach to statistics, and review for the final exam.

References:

1. https://mathworld.wolfram.com/FishersExactTest.html
2. https://online.stat.psu.edu/stat504/lesson/4/4.5
3. https://www.statology.org/fishers-exact-test/
4. https://www.stats.ox.ac.uk/~winkel/stat2l5+1up.pdf
5. https://bookdown.org/anshul302/HE802-MGHIHP-Spring2020/Chi2.html
6. https://api.taylorfrancis.com/content/books/mono/download?identifierName=doi&identifierValue=10.1201/9781420035957&type=googlepdf
7. https://www.reneshbedre.com/blog/fisher-exact-test.html
8. https://search.r-project.org/CRAN/refmans/CNPS/html/permu_table.html