Lecture #16

Recall: Want a way to do hypothesis testing on the parameters of a multinomial distribution $Multi(\vec{p}; n)$ where n (# trials) is known but $\vec{p} = (p_1, \dots, p_k)$ is not.

- So, we would consider hypotheses of the form $H_0: \vec{p} = \vec{p}_0 = (p_1^0, \dots, p_k^0)$
  $H_1: \vec{p} \neq \vec{p}_0$

To understand how to evaluate such hypotheses, we should understand the behavior of a particular observation under the assumption H_0 holds.

- So, suppose $(X_1, \dots, X_k)$ is observed from $Multi(\vec{p}; n)$. We saw that while an individual X_i is asymptotically normal in the sense that

$$\frac{X_i - np_i}{\sqrt{np_i}} \rightsquigarrow N(0, \sqrt{1-p_i}),$$

the global correlation amongst the $\left\{\frac{X_i - np_i}{\sqrt{np_i}}\right\}_{i=1}^{k}$, due to the constraint $\sum_{i=1}^{k} X_i = n$, prevents a simple (approximate/asymptotic) analysis of something like $\sum_{i=1}^{k} \frac{X_i - np_i}{\sqrt{np_i}}$.

- Instead, we saw that a delicate linear algebraic analysis of the joint distribution $\left(\frac{X_1 - np_1}{\sqrt{np_1}}, \dots, \frac{X_k - np_k}{\sqrt{np_k}}\right)$ yields that under H_0,

$$\sum_{i=1}^{k} \frac{[X_i - np_i]^2}{np_i} \rightsquigarrow \chi^2(k-1),$$

where $\chi^2(k-1)$ is a χ^2-distribution with $(k-1)$ degrees of freedom:

$$\chi^2(k-1) = \sum_{i=1}^{k-1} Z_i^2 \quad , \text{ where } \{Z_i\}_{i=1}^{k-1} \text{ are i.i.d. standard normals.}$$

- This gives us a framework for testing if $\sum_{i=1}^{k} \frac{[x_i - np_i]^2}{np_i}$ is anomalous or not: compare it to a cut-off χ_α^2 where

$$\mathbb{P}\left(\chi^2(k-1) > \chi_\alpha^2\right) = \alpha$$

for some significance level α.

- Just like for a Wald Test, we can use a computer/standard table to compute χ_α^2 for fixed (α, k) values.

- <u>To summarize</u>: To run a Pearson's χ^2 test on $H_0: \vec{p} = \vec{p}_0 = (p_1^0, \dots, p_k^0)$
$H_1: \vec{p} \neq \vec{p}_0$,

compute the test statistic $T = \sum_{i=1}^{k} \frac{[x_i - np_i^0]^2}{np_i^0}$. Let α be the specified significance level. We reject H_0 at level α if $T > \chi_\alpha^2$, where χ_α^2 is chosen s.t. $\mathbb{P}(\chi^2(k-1) > \chi_\alpha^2) = \alpha$.

<u>Note</u>: The cut-off χ_α^2 depends on both the significance level (α) and the number of degrees of freedom $(k-1)$.

ex: Suppose we have a 6-sided die. We roll it 6,000 times and observe the following outcome:

Roll	Frequency
1	971
2	1084
3	1013
4	918
5	972
6	1042

Can we argue the die is unfair? Let's phrase "fairness" as a statement about multinomial distribution. $\underline{X} = (\underline{X}_1, \underline{X}_2, \dots, \underline{X}_6) \sim Multi(\vec{p}, 6000)$

$H_0: \vec{p} = (\frac{1}{6}, \frac{1}{6}, \dots, \frac{1}{6})$

$H_1: \vec{p} \neq (\frac{1}{6}, \dots, \frac{1}{6})$.

Well, the χ^2 test statistic is $T = \sum_{i=1}^{k} \frac{[x_i - np_i^0]^2}{np_i^0}$

In our case, $p_i^0 = 1/6$ for all $i = 1, \dots, 6$; this is the quantification of fairness.

Then $n = 6000$, $k = 6$, and we compute

$$T = \sum_{i=1}^{6} \frac{[x_i - 6000 \cdot 1/6]^2}{6000 \cdot 1/6}$$

$$= \frac{[971-1000]^2}{1000} + \frac{[1084-1000]^2}{1000} + \frac{[1013-1000]^2}{1006}$$

$$+ \frac{[918-1000]^2}{1000} + \frac{[972-1000]^2}{1000} + \frac{[1042-1000]^2}{1000}$$

$$\approx 17.3380$$

So, how do we use this to evaluate H_0? Lets run the test at $\alpha = .01$.

(4)

The cut-off χ^2_α associated to $\alpha = .01$ at $6-1=5$ degrees of freedom is (using computer) $\chi^2_\alpha = 15.0863$. Then since $T = 17.3380 > 15.083$, we reject H_0 at the $\alpha = .01$ level.

What is the p-value? Well, what is the value $\mathbb{P}(\chi^2(5) > 17.3380)$
$= .0039$.

So, quite decent evidence the die is "unfair", albeit not very unfair.

<u>Remark</u>: Unlike Gaussian r.v., χ^2 r.v. are not symmetric:

As # d.o.f. increases, the bulk/center of the distribution increases, as does the variance (HW)

- Notice that $\mathbb{P}(\chi^2(k) > 0) = 1 \ \forall k$, but $\mathbb{P}(\chi^2(k) < N) < 1$ for all k, N

In other words, χ^2 distributions have only 1 infinite tail, the one towards $+\infty$.

Recap of Hypothesis Testing: We consider hypotheses around parameters of models,

i.e. $H_0: \mu = \mu_0$ or $H_0: \vec{p} = (p_1^0, \dots, p_k^0)$

$H_1: \mu \neq \mu_0$ $\quad$ $H_1: \vec{p} \neq (p_1^0, \dots, p_k^0)$

- for μ the mean of a Gaussian
- for $\vec{p}$ the probabilities associated to a multinomial distribution

- To evaluate these hypotheses (or any), we needed to answer: "what is the distribution of the sample statistic under the assumption H_0 holds"
- The Wald Test answered that question by leveraging the CLT.
- The Pearson χ^2 test leveraged delicate linear algebra.
- In general, one can be creative and try to develop understanding of the test statistic under H_0; This is part of what research in mathematical statistics looks like!
- Going forward, we will consider more non-parametric/vaguer hypothesis tests.

Our focus will be on tests for independence: are two R.V. X, Y independent?