

# Stat 401, section 14.2 Goodness of Fit – Composite Hypotheses

notes by Tim Pilachowski

From Lecture 14.1: A **multinomial experiment** generalizes a binomial experiment by allowing each trial to result in one of  $k$  possible outcomes, where  $k > 2$ . We referred to the  $k$  possible outcomes on any given trial as categories, and  $p_i$  denoted the probability that a trial results in category  $i$ . That is,  $p_i$  was interpreted as the proportion of the population falling in the  $i^{\text{th}}$  category.

The null hypothesis of interest specified the value of each  $p_i$ :  $H_0: p_1 = p_{10}, p_2 = p_{20}, \dots, p_k = p_{k0}$ .

The alternative hypothesis stated that  $H_0$  was not true – that is, that at least one of the  $p_i$ 's has a value different from that asserted by  $H_0$  (implying that at least two must be different, since all of the proportions added together must equal 1).

In a multinomial experiment the expected number of trials resulting in category  $i$  is

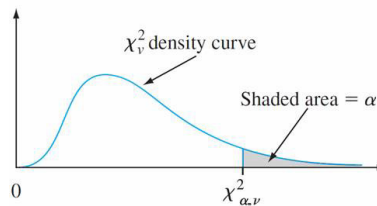
$$E(N_i) = np_i \quad (i = 1, 2, \dots, k).$$

When  $H_0: p_1 = p_{10}, p_2 = p_{20}, \dots, p_k = p_{k0}$  is true, these expected values become

$$E(N_1) = np_{10}, E(N_2) = np_{20}, \dots, E(N_k) = np_{k0}.$$

The test statistic for a multinomial experiment is  $\sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}} = \sum \frac{(n_i - np_{i0})^2}{np_{i0}}$ .

The probability distribution is a **chi-squared ( $\chi^2$ ) distribution** which has  $\nu = k - 1$  degrees of freedom. The critical value,  $\chi^2_{\alpha, \nu}$ , is the value such that  $\alpha$  of the area under the  $\chi^2$  curve lies to the right of  $\chi^2_{\alpha, \nu}$ .



Selected values of  $\chi^2_{\alpha, \nu}$  are given in Appendix Table A.7. A value of the test statistic which is greater than the value of the critical value (calculated  $\chi^2 \geq \chi^2_{\alpha, k-1}$ ) will imply  $P\text{-value} \leq \alpha$ , and we reject the null hypothesis.

The null hypothesis in Lecture 14.1 is a **simple hypothesis** in the sense that each  $p_{i0}$  was a specified number, so that the expected cell counts when  $H_0$  was true were uniquely determined numbers. Now, in Lecture 14.2, we are going to consider one (and only one) application for which each  $p_{i0}$  is determined by some outside parameter: Goodness of Fit for Continuous Distributions, specifically, determining whether a set of data indicates that the population has an underlying normal distribution.

If  $H_0$  states that the underlying distribution is normal, use of a  $\chi^2$  test must be preceded by a choice of cells and estimation of  $\mu$  and  $\sigma$ . To ensure that the chi-squared test is valid, the cells should be chosen independently of the sample observations. Once the cells are chosen, it is almost always quite difficult to estimate unspecified parameters (such as  $\mu$  and  $\sigma$  in the normal case) from the observed cell counts, so instead maximum likelihood estimates (see Lecture 6.2) based on the full sample are computed.

Theorem: Let  $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m$  be the maximum likelihood estimators of  $\theta_1, \theta_2, \dots, \theta_m$  based on the full sample  $X_1, X_2, \dots, X_n$ , and let  $\chi^2$  denote the statistic based on these estimators. Also let  $P_1 =$  the  $P$ -value for an upper-tailed chi-squared test based on  $k - 1$  degrees of freedom, and let  $P_2 =$  the  $P$ -value for an upper-tailed chi-squared test based on  $k - 1 - m$  degrees of freedom. Then it can be shown that  $P_1 \leq P\text{-value} \leq P_2$ .

For a given significance level  $\alpha$ : 1) If  $\alpha \leq P_1$ , fail to reject  $H_0$ ; If  $\alpha \geq P_2$ , reject  $H_0$ ; If  $P_1 < \alpha < P_2$ , withhold judgement.

We're going to illustrate the process by Example.

Example A: The SAT is designed so that the observed scores will have normal distribution. Do the 162 scores listed below support this hypothesis? ( $\alpha = 0.05$ )

800	800	790	780	760	760	760	760	750	740	740	740	740	740	740
730	720	720	720	720	720	710	710	710	700	700	700	700	700	700
700	700	700	700	700	690	690	690	690	690	690	690	690	680	680
680	680	680	680	680	680	670	670	670	670	660	660	660	650	650
650	650	650	650	650	650	650	650	640	640	640	640	640	640	640
640	640	640	630	630	630	630	630	630	630	630	620	620	620	620
610	610	610	610	600	600	600	600	590	590	590	580	580	580	570
570	570	570	570	570	570	570	560	560	560	560	560	560	550	550
550	540	540	540	540	540	540	540	530	530	530	530	530	530	520
520	520	520	510	510	500	500	490	490	480	470	460	450	440	440
430	430	420	410	410	410	410	400	400	390	390	360			

hypotheses:

Prior to sampling, it is presumed that the population mean  $\mu$  is 600 with standard deviation  $\sigma$  of 100.

As noted at the end of Lecture 14.1b, the eight  $z$ -intervals  $(-\infty, -1.15)$ ,  $[-1.15, -0.675)$ ,  $[-0.675, -0.32)$ ,  $[-0.32, 0)$ ,  $[0, 0.32)$ ,  $[0.32, 0.675)$ ,  $[0.675, 1.15)$  and  $[1.15, \infty)$  have a probability of  $1/8$  each, i.e. the probability is uniform for each interval.

For  $\mu = 600$  and  $\sigma = 100$ , these intervals transform to  $X$  boundaries, (Fill in "Cell" column below.)

The observed counts are ... (Fill in "Observed" column below.)

Next step: We need to transform these  $X$  boundaries into estimated (expected) cell probabilities,  $\pi_i(\hat{\mu}, \hat{\sigma})$ , using the maximum likelihood estimates  $\hat{\mu}$  and  $\hat{\sigma}$ . Back in Lecture 6.2, we derived formulas for both of these.

Last step before calculating the  $\chi^2$  statistic: computing estimated expected cell counts,  $n * \pi_i(\hat{\mu}, \hat{\sigma})$ . (Fill in "Estimated Expected" column below.)

Cell	Observed $n_i$	Estimated Expected $= n\pi_i(\hat{\mu}, \hat{\sigma})$	$(O - E)$	$(O - E)^2$	$\frac{(O - E)^2}{E} = \frac{(n_i - n\pi_i(\hat{\mu}, \hat{\sigma}))^2}{n\pi_i(\hat{\mu}, \hat{\sigma})}$
				$\chi^2 = \sum \frac{(O-E)^2}{E}$	

**IMPORTANT:** The symbol  $\chi^2$  is a *notation!* Do not square the sum in the last column.

critical values:

conclusion: