STATS 305B: Methods for Applied Statistics I **Winter 2016/17**

Lecture 2: January 11

Lecturer: Jonathan Taylor *Scribes: Kenneth Tay*

2.1 (Agresti 1.3, 1.4) Inference for Binomials

Suppose we want to infer the binomial proportion π based on observing $Y \sim \text{Binom}(n, \pi)$. The maximum likelihood estimator is simply

$$\hat{\pi}_{MLE} = \hat{\pi} = \frac{Y}{n}.$$

Let's say now that we wish to either test $H_0 : \pi = \pi_0$, or find a 95% confidence interval for π. There are 3 possible approaches. All of them will use the Fisher information "matrix":

$$I(\pi) = \mathbb{E}_\pi \left[-\frac{\partial^2}{d\pi^2} \log L(\pi \mid Y) \right] = \frac{n}{\pi(1-\pi)}.$$

2.1.1 Wald Test

$\hat{\pi} \approx \mathcal{N}\left(\pi_0, (I(\hat{\pi})^{-1})\right)$, so the corresponding z-statistic is

$$z_{Wald} = \frac{\hat{\pi} - \pi_0}{\sqrt{\hat{\pi}(1-\hat{\pi})/n}}.$$

The 95% confidence interval is given by

$$\left(\hat{\pi} - 1.96\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}, \hat{\pi} + 1.96\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}} \right).$$

As these statistics/intervals are easy to calculate, they are what software usually return.

2.1.2 Score Test

Let $S = \left.\frac{\partial L(\pi \mid Y)}{\partial \pi}\right|_{\pi_0}$. Under H_0, $S \approx \mathcal{N}(0, I(\pi_0))$, and the corresponding z-statistic is

$$z_{score} = \frac{\hat{\pi} - \pi_0}{\sqrt{\pi_0(1-\pi_0)/n}}.$$

The 95% confidence interval is given by

$$\left\{ \pi : \left| \frac{\hat{\pi} - \pi}{\sqrt{\pi(1-\pi)/n}} \right| \le 1.96 \right\}.$$

2.1.3 Likelihood-Ratio Test

The binomial log-likelihood function is equal to $L_0 = y \log \pi_0 + (n-y)\log(1-\pi_0)$ under H_0, and is equal to $L_1 = y \log \hat{\pi} + (n-y)\log(1-\hat{\pi})$ more generally. Hence, the likelihood-ratio test statistic simplifies:

$$-2(L_0 - L_1) = 2\left[y \log \frac{\hat{\pi}}{\pi_0} + (n-y)\log\frac{1-\hat{\pi}}{1-\pi_0}\right],$$

which has approximate distribution χ_1^2 under H_0.

The 95% confidence interval is given by

$$\left\{\pi : -2(L_0 - L_1) \le \chi^2_{1,1-\alpha} = (1.96)^2\right\}.$$

2.2 (Agresti 1.5) Inference for Multinomials

Suppose $Y \sim \text{Multinom}(n, \pi)$, where $\pi \in \mathbb{R}^k$ represents the probability of Y being in each of k categories.

If we let $\Sigma(\pi) = \text{diag}(\pi) - \pi\pi^T$, then we have $\text{Cov}(Y) = n\Sigma(\pi)$. (Note: The rank of $\Sigma(\pi)$ is $k-1$.) As $n \to \infty$, $\frac{Y}{n} - \pi = \hat{\pi}_{MLE} - \pi \sim \mathcal{N}\left(0, \frac{\Sigma(\pi)}{n}\right)$.

To test $H_0 : \pi = \pi_0$, there are 2 commonly used goodness-of-fit tests:

2.2.1 Likelihood-Ratio Test

Under H_0, we have the likelihood-ratio test statistic

$$G(\pi_0, \hat{\pi}) = 2\sum_{j=1}^{k} y_j \log\left(\frac{y_j/n}{\pi_{0,j}}\right) \approx \chi^2_{k-1}.$$

2.2.2 Pearson's χ^2

We have the test statistic

$$\sum_{j=1}^{k} \frac{(O_j - E_j)^2}{E_j} = \sum_{j=1}^{k} \frac{(y_j - n\pi_{0,j})^2}{n\pi_{0,j}} \approx \chi^2_{k-1}$$

under H_0. (In the above, O_j and E_j refer to the observed and expected values of Y_j.) This test has origins in the Mahalanobis distance test for the Gaussian.

We end of this lecture with a fact for normals:

Theorem 2.1 *Suppose $Z \sim \mathcal{N}(\mu, \Sigma)$. Then*

$$(Z-\mu)^T \Sigma^{\dagger}(Z-\mu) \sim \chi^2_{rank(\Sigma)},$$

where $\Sigma^{\dagger}$ is the pseudo-inverse of Σ.