Sections 1.3-1.5

Statistical Inference
Objectives

• By the end of this presentation, we expect that you will

  – Learn some of the techniques used in statistical inference

  – And understand significance tests and confidence intervals for binomial and multinomial parameters.
Inferences about parameters

- Likelihood: $prob(Data/Parameter)$

- The maximum likelihood estimator ($\hat{\pi}$) is that value of $\pi$ that maximizes the likelihood for a given data

- An estimator is random whereas an estimate is a fixed
Properties of MLEs

• They have large sample normal distributions:

• They are asymptotically consistent: \( \hat{\pi} \to \pi \) as the sample size \( n \to \infty \).

• They are asymptotically efficient: \( var(\hat{\pi}) \to 0 \) as the sample size \( n \to \infty \).

• useful for constructing hypothesis tests and confidence intervals

• Under \( H_0 \), the Wald, Score and likelihood ratio test statistics are approximately \( \chi^2 \)
Wald, likelihood ratio, and score tests

- Consider an example of the significance test for binomial parameter

\[ H_0: \pi = \pi_0 \]
\[ H_0: \pi \neq \pi_0 \text{ (or 1-sided)} \]

- Wald statistics:

\[ z = \frac{p - \pi_0}{SE} = \frac{p - \pi_0}{\sqrt{\hat{\pi}(1-\hat{\pi})/n}} \]

- Score test:

\[ z = \frac{p - \pi_0}{SE} = \frac{p - \pi_0}{\sqrt{\pi_0(1-\pi_0)/n}} \]

- Likelihood ratio:

\[ \Lambda = \frac{\text{maximum under reduced model} \ (H_0)}{\text{maximum full model}} \]
Confidence Interval

- Let’s consider $H_0 : \pi = \pi_0$ where $\pi_0$ is fixed and known (e.g. $H_0 : \pi = 0.5$).

- The Wald test plugs in the MLE $\hat{\pi} = \frac{y}{n}$ for the unknown $\pi$ in the large sample SE: that is $se(\hat{\pi}) = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}$

- So that a 95% CI for $\pi$ is $\hat{\pi} \pm 1.96 \times SE(\hat{\pi}) = \hat{\pi} \pm 1.96 \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}$

- The score test plugs in null value into the large sample SE: that is $SE(\pi_0) = \sqrt{\frac{\pi_0(1-\pi_0)}{n}}$

- So that a 95% CI for $\pi$ is $\pi_0 \pm 1.96 \sqrt{\frac{\pi_0(1-\pi_0)}{n}}$
Likelihood function and ML estimate for binomial parameter

Note:

• Wald CI often has poor performance in categorical data analysis unless \( n \) quite large.

• Wald denominator uses estimated SE whereas the score SE is assumed known

• For inference about proportions, score method tends to perform better than Wald method, in terms of having actual error rates closer to the advertised levels.

• For small \( n \), inference uses actual binomial sampling dist. of data instead of normal approx. for that dist.
Confidence Interval: Vegetarian example

In a study of 25 students, none (y=0) were vegetarians. What’s the 95% CI for the proportion of the vegetarians?

• Wald CI gives (0,0)

• Score CI gives (0.0, 0.133)

• LR CI gives (0.0, 0.074)

• When $\pi \sim 0$, the sampling distribution of the estimator is highly skewed to the right. Worth considering alternative methods (e.g. exact methods) not requiring asymptotic approximations.