Uncertainty in estimation

Wenda Zhou

June 2, 2017

Uncertainty in estimation

We have seen how to compute a single estimate $\hat{\theta}$ of a parameter.
Can we also quantify the uncertainty in the estimate?

Uncertainty in estimation

Example: coin toss

Suppose X follows a binomial distribution with parameters n and p. We have seen the m.l.e. estimator for p given by:

$$\hat{p} = \frac{1}{n} X \tag{1}$$

Uncertainty in estimation

Example: coin toss

Suppose X follows a binomial distribution with parameters n and p. We have seen the m.l.e. estimator for p given by:

$$\hat{p} = \frac{1}{n}X \tag{1}$$

Intuition: estimator should be better for $n = 100$ than $n = 10$.

Uncertainty in estimation

Example: coin toss

Recall: for unbiased estimators, lower variance is better.
Let us compute:

$$\mathrm{Var}_p \hat{p} = \frac{1}{n} p(1-p) \tag{2}$$

Uncertainty in estimation

Example: coin toss

Recall: for unbiased estimators, lower variance is better.
Let us compute:

$$\mathrm{Var}_p\, \hat{p} = \frac{1}{n} p(1-p) \tag{2}$$

Larger n leads to lower variance and better estimation. Hence may also wish to report standard error.

Confidence intervals

Instead of giving a single point as estimate: give range of values.

Confidence interval

A confidence interval is an interval that contains the true parameter with high probability.

Confidence intervals

Instead of giving a single point as estimate: give range of values.

Confidence interval

A confidence interval is an interval that contains the true parameter with high probability.

Note: the interval depends on the data, thus is random. The parameter is fixed (although unknown)

Confidence intervals

Example: toin coss

Suppose that we toss a coin 100 times, and observe exactly 50 heads.

The computed 95% confidence interval is: $[0.398, 0.602]$.

Confidence intervals

Example: toin coss

Suppose that we toss a coin 100 times, and observe exactly 50 heads.
The computed 95% confidence interval is: $[0.398, 0.602]$.
Suppose instead that we toss a coin 1000 times, and observe exactly 500 heads.
The computed 95% confidence interval is: $[0.469, 0.531]$.

Confidence intervals

Example: toin coss

Suppose that we toss a coin 100 times, and observe exactly 50 heads.

The computed 95% confidence interval is: $[0.398, 0.602]$.

Suppose instead that we toss a coin 1000 times, and observe exactly 500 heads.

The computed 95% confidence interval is: $[0.469, 0.531]$.

The second confidence interval is narrower, reflecting our increased certainty.

Confidence intervals

Interpreting the 95%

Let us write the confidence interval as $[a(X), b(X)]$.
Then, $[a(X), b(X)]$ is a 95% confidence interval if:

$$\mathrm{P}_p(a(X) \leq p \leq b(X)) = 0.95 \tag{3}$$

Confidence intervals

Interpreting the 95%

Let us write the confidence interval as $[a(X), b(X)]$.
Then, $[a(X), b(X)]$ is a 95% confidence interval if:

$$\mathrm{P}_p(a(X) \leq p \leq b(X)) = 0.95 \tag{3}$$

If we repeat the experiment and compute a confidence interval in the same manner, it will cover the true value 95% of the time on average.

Computing a confidence interval

Normal distribution

Suppose that we observe a single observation:

$$\hat{\theta} \sim \mathcal{N}(\theta, \sigma_\theta^2) \tag{4}$$

σ_θ is the standard error (known).
Let us compute a confidence interval for θ.

Computing a confidence interval

Normal distribution

A $(1-\alpha)$ confidence interval is given by:

$$[\hat{\theta} - \sigma_\theta z_{1-\alpha/2}, \hat{\theta} + \sigma_\theta z_{1-\alpha/2}] \tag{5}$$

Where z_β is such that $\mathrm{P}(Z \leq z_\beta) = \beta$ (quantile).

Computing a confidence interval

Example

Suppose that after running an experiment, we obtain an estimate of $\hat{\theta} = 4.5$ and standard error 1.2. Compute a 95% estimate for θ.
We have that: $\alpha = 0.025$, hence $z_{1-\alpha/2} = 1.96$.
Hence we have that a 95% confidence interval for this value is:

$$[4.5 - 1.96 \times 1.2, 4.5 + 1.96 \times 1.2] = [2.15, 6.85] \qquad (6)$$