Math 3215: Lecture 17

Will Perkins

March 13, 2012

1 A Proof of the CLT

def. The *characteristic function* of a random variable X is

$$\phi_X(t) := \mathbb{E}e^{itX}$$

Let $X \sim N(0,1)$. Then $\phi_X(t) = e^{-t^2/2}$.

Steps in the proof of the CLT:

1. Know the power series expansion $e^x = 1 + x + x^2/2 + x^3/3! + \dots$
2. Know the limit $(1 + y/n)^n \to e^y$
3. Let $X_1, \dots X_n$ be iid w characteristic function $\phi_X(t)$, mean 0 and variance σ^2. Let $Y_n = \frac{X_1 + \dots + X_n}{\sigma\sqrt{n}}$.
4. What is $\phi_{Y_n}(t)$?
5. Fix t, take limit: $\lim_{n\to\infty} \phi_{Y_n}(t)$. What do you get?

2 Parameter Estimation

When we take a poll of potential voters before the presidential election, what do we really want to know?

- A yes/no question: do more voters prefer candidate A than candidate B?
- What's a good estimate for the fraction of voters who prefer candidate A?
- How accurate is that estimate?

Estimating the fraction of voters who prefer a particular candidate is called making a *point estimate*. Other examples of point estimation:

- Estimate the average height of a Georgia Tech student
- Estimate the chance that a new medicine will cure a cold
- Estimate the bias of a roulette wheel
- Estimate the error rate of a manufacturing process
- Estimate the average income of a resident of Georgia

In all of the above exmaples we are looking for a particular number, a *parameter* of the unknown distribution. We often assume the underlying distribution has a particular form (Binomial, Coin Flip, Poisson, Geometric, Normal) but we don't know the parameter. We then observe samples from the distribution and then make an estimate.

3 A Basic Example

Politcal polling or coin flipping is the most basic example of parameter estimation. There's an unknown distribution X which takes value 1 with probability p and 0 with probability $1-p$. We observe n samples, $X_1, X_2, \ldots X_n$, iid copies of X and want to estimate p. What's a good estimate? Let $x_1, \ldots x_n$ be the observed values of $X_1, \ldots X_n$

The first idea that comes into your head is right: let $\hat{\Theta} = \frac{x_1+\cdots+x_n}{n}$ be the sample mean. This is the most natural estimate for p.

4 Maximum Likelihood Estimator

How can we show that $\hat{\Theta}$ is the best estimate mathematically?

We can compute the probability of seeing the sample $x_1, \ldots x_n$ given that $p = \Theta$:

$$L(\Theta) = \Theta^{\sum x_i}(1-\Theta)^{n-\sum x_i}$$

One way to find the best estimate for p is to find the value of Θ that maximizes $L(\Theta)$. This is called the *maximum likelihood estimator*, and we denote it $\hat{\Theta}$.

- Using basic calculus, find $\hat{\Theta}$
- $\hat{\Theta}$ is a random variable. Its value depends on the particular values $x_1, \ldots x_n$.
- What is $\mathbb{E}\hat{\Theta}$?
- What is $var(\hat{\Theta})$?
- Using the CLT what is the probability (depending on n) that $\hat{\Theta}$ is within .05 of p?
- within .01?

4.1

Some values of $\Phi(t)$:

- $\Phi(1) - \Phi(0) = .34$
- $\Phi(2) - \Phi(0) = .477$
- $\Phi(3) - \Phi(0) = .4987$