

Stat 324: Lecture 13

Sampling distributions

Moo K. Chung
mchung@stat.wisc.edu

March 10, 2005

Concepts

1. Random variables $X_1, \dots, X_n$ form a *random sample* if they are independent and identically distributed (i.i.d.). A *statistic* is a function of the random sample, i.e. $h(X_1, \dots, X_n)$. The distribution of the statistic is called the *sampling distribution*.
2. Suppose $\mathbb{E}X_i = \mu, \mathbb{V}X_i = \sigma^2$. Then $\mathbb{E}\bar{X} = \mu, \mathbb{V}\bar{X} = \sigma^2/n$. This can be shown from the fact that $\mathbb{E}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathbb{E}X_i, \mathbb{V}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathbb{V}X_i$.
3. If a random sample $X_i \sim N(\mu, \sigma^2)$, then $\bar{X} \sim N(\mu, \sigma^2/n)$.
4. *Central limit theorem*. Consider a random sample $X_1, \dots, X_n$ with mean μ and variance σ^2 . For sufficiently large n , $\bar{X} \sim N(\mu, \sigma^2/n)$. The larger the value of n , the better the approximation. Similarly $Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$.

Problems

1. Let $X_1, \dots, X_n \stackrel{i.i.d.}{\sim} N(\mu, \sigma^2)$. Then are $\bar{X}$ and $X_j - \bar{X}$ independent?

solution. $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$. Since $X_j - \bar{X}$ is a linear combination of X_i 's, it is also distributed as normal. Any uncorrelated normal distributions are independent. So we only need to check if $\text{Cov}(\bar{X}, X_j - \bar{X}) = 0$.

$$\begin{aligned}\text{Cov}(\bar{X}, X_j - \bar{X}) &= \text{Cov}\left(\frac{1}{n} \sum_{i=1}^n X_i, X_j - \frac{1}{n} \sum_{k=1}^n X_k\right) \\ &= \frac{1}{n} \sum_{i=1}^n \text{Cov}(X_i, X_j) - \frac{1}{n^2} \sum_{i,k=1}^n \text{Cov}(X_i, X_k) = \frac{1}{n} \mathbb{V}(X_j) - \frac{1}{n^2} \sum_{i=1}^n \mathbb{V}(X_i) \\ &= \frac{1}{n} \mathbb{V}(X_j) - \frac{1}{n} \mathbb{V}(X_j) = 0\end{aligned}$$

2. The service times for customers coming through a checkout counter in a retail store are independent random variables with mean 2 minutes and variance 5. (Hint: Apply the central limit theorem to answer the following questions)
 - (a) Approximate the probability that 100 customers can be served in less than 3 hours of total service time.
 - (b) Find the maximum number of customers that can be served in less than 3 hours with probability 0.9.
 - (c) Approximate the probability that 100 customers can be served in between 3 hours and 3 hours 40 minutes.

Assigned Problems: Exercise 5.48. Exercise 5.50. Exercise 5.90. **Homework 4** due on March 17th. 9:30AM. Solve assigned problems in Lecture 11-14. Please read Chapter 6.