Lecture 17.

From last time:

- A statistic is any function of a random variable
- A statistic $h(X_1, \dots, X_n)$ is itself a random variable, and so has some probability distribution.
- Such a distribution is called a sampling distribution.

  ↳ ex: $\overline{X} = \frac{1}{n}(X_1 + \cdots + X_n)$

  is a sampling distribution for the mean of X.

- A sampling distribution for a RV X depends on the sample size, sample method, distribution of X and other factors.

However, we have the following useful fact:

Central Limit Theorem: Suppose $X_1, \dots, X_n$ is a random sample taken from a population distributed with mean μ and variance σ^2 (i.e. $X_1, \dots, X_n$ are indep. identically distributed). Then, if $\overline{X}_n = \frac{1}{n}(X_1 + \dots + X_n)$, we have

$$\lim_{n \to \infty} \frac{\overline{X}_n - \mu}{\sigma / \sqrt{n}} = \underbrace{N(0,1)}_{\text{standard normal.}}$$

This is statistically very useful, since it implies that for any random variable X with mean μ and variance σ^2, the sampling distribution

$$\overline{X} = \frac{1}{n}(X_1 + \dots + X_n)$$

can be approximated by the normal random variable $N(\mu, \sigma^2/n)$.

In practice taking $n \geq 30$ is often (not always) enough to see the effects of the central limit theorem.

For example: If we roll a fair die and set X = the value of the roll, then X is a uniform discrete RV with $\mu = E[X] = 3.5$ and $\sigma^2 \approx 2.9167$.

If we roll, say 30 dice, and set $\overline{X}$ to be the average of the rolls then $\overline{X}$ is still a discrete random variable, but it is well approximated by the continuous normal random variable $N(3.5, \frac{\sigma^2}{30}) = N(3.5, 0.972)$.

- whether the central limit theorem applies for small n depends on the original distribution of X. If X has a symmetric, unimodal distribution, then $\overline{X}$ can be approximated by a normal distribution with as few as 4 or 5 samples.

Ex: An electronic company manufactures resistors that have a mean resistance of 100Ω and a standard deviation of 10Ω. If the resistance is normally distributed, what is the the probability that a sample of 25 has an average resistance of ≤ 95?

Solution: $\overline{X}$, the sampling distribution of the mean is normally distributed with mean $\mu = 100$ and standard deviation $\sigma = \frac{10}{\sqrt{25}} = \frac{10}{5} = 2$.

Thus, we want to find

$$P(\overline{X} < 95) = P\left(Z < \frac{95-100}{2}\right)$$
$$= P(Z < -2.5) = \Phi(-2.5)$$
$$\approx 0.0062$$

Remark: In practical application, this means that a random sample of resistors with mean resistance less than 95 is rare. If you did find a sample with mean resistance < 25, you should doubt whether the mean and std. dev is actually 100 and 10.

General Concepts of Point Estimation

(7.3.1) Unbiased Estimators.

- Recall that a parameter θ is some numerical feature/property of a population/data set.
- An estimator $\hat{\Theta}$ for θ is a statistic (function of a random sample) used to estimate θ.

Defn: We define the bias of an estimator $\hat{\Theta}$ to be the quantity

$$E(\hat{\Theta}) - \theta.$$

We say that $\hat{\Theta}$ is an unbiased estimator if $E(\hat{\theta}) - \theta = 0$. (so $E(\hat{\theta}) = \theta$.

Ex: Suppose $X_1, \dots X_n$ is a random sample from a population represented by a random variable X, with mean μ and variance σ^2.

We know that

$$E[X_1 + \dots + X_n] = E[X_1] + \dots + E[X_n]$$
$$= \mu + \dots + \mu = n\mu.$$

Hence $E[\overline{X}] = E\left[\frac{1}{n}(X_1 + \dots + X_n)\right] = \frac{1}{n}(n\mu) = \mu$.

and so $\overline{X}$ is an unbiased estimator of μ.

Claim: The sample variance S^2 is an unbiased estimator for the variance, σ^2. Why?

$$E(S^2) = E\left[\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}\right] = \underbrace{E\left[\frac{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2}{n-1}\right]}_{\text{shortcut for } s^2.}$$

$$= \frac{1}{n-1}\left(\sum_{i=1}^{n} E[X_i^2] - nE[\overline{X}^2]\right)$$

Now, $E[X_i^2] = \mu^2 + \sigma^2$ and $E[\overline{X}^2] = \mu^2 + \frac{\sigma^2}{n}$

(since $Var(X_i) = E[X_i^2] - \underbrace{E[X_i]^2}_{=\mu^2}$) (similarly) and so we have

$$E(S^2) = \frac{1}{n-1}\left(\sum_{i=1}^{n}(\mu^2+\sigma^2) - n\left(\mu^2+\frac{\sigma^2}{n}\right)\right)$$

$$= \frac{1}{n-1}\left(n(\mu^2+\sigma^2) - n\mu^2 - \sigma^2\right).$$

$$= \frac{1}{n-1}\left(\cancel{n\mu^2} + \overset{n-1}{\cancel{n}}\sigma^2 - \cancel{n\mu^2} - \cancel{\sigma^2}\right) =$$

$$= \frac{1}{n-1}\left((n-1)\sigma^2\right) = \sigma^2.$$

Note, however, that although S^2 is an unbiased estimator for σ^2, S is not an unbiased estimator for σ.

↳ fortunately, as the sample size grows $E[S]-\sigma$ gets smaller so somehow we can use S as a good estimator of σ (next chapter).