Sampling Distribution and Estimators

The value of a statistic, such as the sample mean, depends on the particular values included in the sample, and it generally varies from sample to sample. This variability of a statistic is called sampling variability.

The sampling distribution of mean is the probability distribution of sample means, with all samples having the same sample size $n$.

Examples

P214 Table 5-2 5-3

Tow dice
**Central Limit Theorem**

Under some assumptions, the distribution of sample means $\bar{x}$ will be normally distributed with mean $\mu$ and SD $\frac{\sigma}{\sqrt{n}}$, REGARDLESS of the distribution of the individual observation.

(Approximation of binomial distribution, Poisson distribution, etc.)

---

**Sampling distribution:**

1. Distribution of sample means from different samples of the same size collected from the same population.
2. No matter what is the distribution of the individual from the population, the sampling distribution will always be normal distribution, as long as the sample size is sufficiently large.
3. The above is guaranteed by Central Limit Theorem (CLT), which also provide the mean and variance of the normal distribution, $\left( \mu, \frac{\sigma^2}{n} \right)$.

How large a sample size is considered as sufficiently large?

**Empirical Rules:**

Two magic numbers: 30 and 5
Applying CLT

When working with individual value from a normally distributed population, use the methods of Section 5-3:

\[ Z = \frac{X - \mu}{\sigma} \]

When working with a mean for some sample (or group), (no matter what is the individual distribution), Be sure to use the std. dev. of the sample means instead of std. dev. of the sample or of the population.

\[ Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \]

EX: Ski Gondola Safety

Maximum capacity is 12 people 2004 pounds

Men have average weight of 172 pounds and std. dev. 29 pounds

Find \( P(\text{a randomly selected man is more than 167 pounds}) \)

Find \( P(12 \text{ randomly selected men have mean weight that greater than 167 pounds}) \)
**Approximation to Binomial**

Review Binomial formula for mean and SD: (P 178 or 232)

Why can we do such approximation?

1. CLT and binomial Random variable is actually sum of 0 and 1’s (Bernoulli random variables), which is n times the average of these 0 and 1’s.

2. Another sampling distribution example: Figure 5-18

Ex: At least 120 male in randomly selected 200 people.
0.5 Correction: