Normal Probability Distributions

Glossary

1. Standard Normal Distribution
    a. Bell Shaped
    b. mean = 0
    c. std dev = 1
2. **Normal Distribution equation for continuous random variable**

FORMULA 6-1

$$y = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$$

3. **Continuous Probability Distribution**
    a. area under CPD is 1
    b. Area under the graph is equal to probability
    c. Also called Density Graph

d. Area = Prob = Height * Width

4. **Z-score**
   Distance along the horizontal scale of the standard normal distribution (corresponding to the number of standard deviations above or below the mean);

5. The area corresponding to the region between two z scores can be found by finding the difference between the two areas found in Table A-2.

6. **Critical Value**

   For the standard normal distribution, a critical value is a z score on the borderline separating those z scores that are significantly low or significantly high.

   The expression za denotes the z score with an area of a to its right. (a is the Greek letter alpha.)

   CAUTION When finding a value of za for a particular value of a, note that a is the area to the right of za, but Table A-2 and some technologies give cumulative areas to the left of a given z score. To find the value of za, resolve that conflict by using the value of 1 - a. For example, to find z0.1, refer to the z score with an area of 0.9 to its left.

7. **Non-standard Normal Distribution**

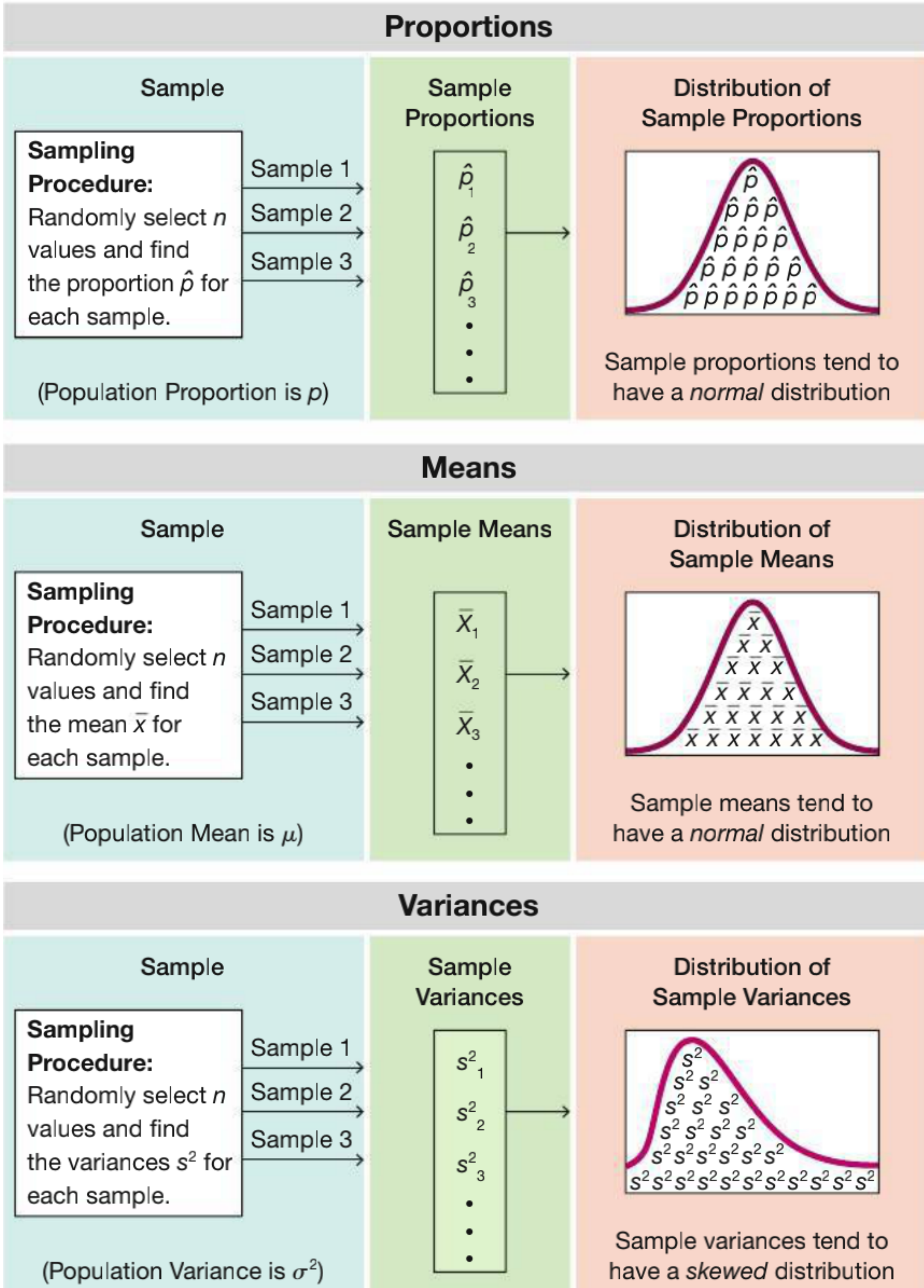
   To standardize x-values to z-scores we use : z=(x-mean)/std dev
   To calculate prob for non-standard distribution
   a. calculate z-score by using formula

   b. use statdisk for area = prob for that z-score.

   To find x-value from area / prob
   a. find z-score in statdisk using cumulative left area.
   b. once we get z-score use this formula to get x-value: x=mean + (z-score*std dev)

   c. or use excel NORMINV { inputs: probability, mean, and standard deviation }

8. **Sampling Distributions and Estimators**



The sampling distribution of a statistic (such as a sample proportion or sample mean) is the distribution of all values of the statistic when all possible samples of the same size n are taken from the same population.

The sampling distribution of the sample proportion is the distribution of sample proportions (or the distribution of the variable pn), with all samples having the same sample size n taken from the same population.

The sampling distribution of the sample mean is the distribution of all possible sample means (or the distribution of the variable x), with all samples having the same sample size n taken from the same population.

The sampling distribution of the sample variance is the distribution of sample variances (the variable s2), with all samples having the same sample size n taken from the same population.

Sample Proportion → Normal Distribution
Sample Mean → Normal Distribution
Sample variance → Skewed Right Distribution

Estimator: An estimator is a statistic used to infer (or estimate) the value of a population
parameter.

Unbiased Estimator: An unbiased estimator is a statistic that targets the value of the corresponding population parameter in the sense that the sampling distribution of the statistic has a mean that is equal to the corresponding population parameter.

Biased Estimators: median, range, standard deviation
Unbiased Estimators: Proportions, mean, variances

9. **Central Limit Theorem**

CENTRAL LIMIT THEOREM

For all samples of the same size n with $n > 30$, the sampling distribution of $\overline{x}$ can be approximated by a normal distribution with mean μ and standard deviation $\sigma/\sqrt{n}$.

The Central Limit Theorem and the Sampling Distribution of $\bar{x}$

Given

1. Population (with any distribution) has mean μ and standard deviation σ.
2. Simple random samples all of the same size n are selected from the population.

Practical Rules for Real Applications Involving a Sample Mean $\bar{x}$

Requirements: Population has a normal distribution *or* $n > 30$:

Mean of all values of $\bar{x}$: $\mu_{\bar{x}} = \mu$

Standard deviation of all values of $\bar{x}$: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

z score conversion of $\bar{x}$: $z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$

Original population is *not* normally distributed *and* $n \le 30$: The distribution of $\bar{x}$ cannot be approximated well by a normal distribution, and the methods of this section do not apply. Use other methods, such as nonparametric methods (Chapter 13) or bootstrapping methods (Section 7-4).

Considerations for Practical Problem Solving

1. **Check Requirements:** When working with the mean from a sample, verify that the normal distribution can be used by confirming that the original population has a normal distribution or the sample size is $n > 30$.
2. **Individual Value or Mean from a Sample?** Determine whether you are using a normal distribution with a *single* value x or the mean $\bar{x}$ from a sample of n values. See the following.
   - Individual value: When working with an *individual* value from a normally distributed population, use the methods of Section 6-2 with $z = \frac{x - \mu}{\sigma}$.
   - Mean from a sample of values: When working with a mean for some *sample* of n values, be sure to use the value of $\sigma/\sqrt{n}$ for the standard deviation of the sample means, so use $z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$.

Correction for a Finite Population

In applying the central limit theorem, our use of $\sigma_{\bar{x}} = \sigma/\sqrt{n}$ assumes that the population has infinitely many members. When we sample with replacement, the population is effectively infinite. When sampling without replacement from a finite population, we may need to adjust $\sigma_{\bar{x}}$. Here is a common rule:

When sampling without replacement and the sample size n is greater than 5% of the finite population size N (that is, $n > 0.05N$), adjust the standard deviation of sample means $\sigma_{\bar{x}}$ by multiplying it by this *finite population correction factor:*

$$\sqrt{\frac{N-n}{N-1}}$$

Except for Exercise 21 "Correcting for a Finite Population," the examples and exercises in this section assume that the finite population correction factor does *not* apply, because we are sampling with replacement, or the population is infinite, or the sample size doesn't exceed 5% of the population size.

10. **Accessing Normality**

    Normal Distribution: The population distribution is normal if the pattern of the points is reasonably close to a straight line and the points do not show some systematic pattern that is not a straight-line pattern.

    Not a Normal Distribution: The population distribution is not normal if either or both of these two conditions applies:

    • The points do not lie reasonably close to a straight line.

    • The points show some systematic pattern that is not a straight-line pattern.

11. **Normal Distribution as an Approximation to the Binomial Distribution**

Normal Distribution as an Approximation to the Binomial Distribution

Requirements

1. The sample is a simple random sample of size n from a population in which the proportion of successes is p, or the sample is the result of conducting n independent trials of a binomial experiment in which the probability of success is p.
2. $np \geq 5$ and $nq \geq 5$.

(The requirements of $np \geq 5$ and $nq \geq 5$ are common, but some recommend using 10 instead of 5.)

Normal Approximation

If the above requirements are satisfied, then the probability distribution of the random variable x can be approximated by a normal distribution with these parameters:

- $\mu = np$
- $\sigma = \sqrt{npq}$

Continuity Correction

When using the normal approximation, adjust the discrete whole number x by using a *continuity correction* so that any individual value x is represented in the normal distribution by the interval from $x - 0.5$ to $x + 0.5$.

Procedure for Using a Normal Distribution to Approximate a Binomial Distribution

1. Check the requirements that $np \geq 5$ and $nq \geq 5$.
2. Find $\mu = np$ and $\sigma = \sqrt{npq}$ to be used for the normal distribution.
3. Identify the discrete whole number x that is relevant to the binomial probability problem being considered, and represent that value by the region bounded by $x - 0.5$ and $x + 0.5$.
4. Graph the normal distribution and shade the desired area bounded by $x - 0.5$ or $x + 0.5$ as appropriate.