Sample Size and Statistics

Dr. Richard Jerz

A Marketing Example

How to Calculate Sample Size

$$n = \frac{z^2(pq)}{e^2}$$

Where
- n = the calculated sample size
- z = standard error associated with the chosen level of confidence (typically, 1.96)
- p = estimated percentage in the population
- q = $(100\% - p)$
- e = acceptable error (desired accuracy level)

Equations for Sample Size

$$n = p(1-p)\left(\frac{z}{E}\right)^2 \qquad n = \left(\frac{z\sigma}{E}\right)^2$$

Reference p316-317

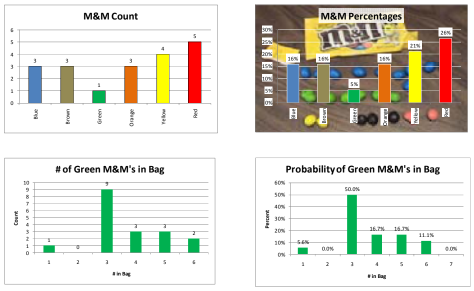
Statistical Concepts

- Descriptive versus inferential statistics
- Population versus sample
- Discrete versus continuous data
- Probability distributions
- Probability concepts
- Binomial, Normal, and student-t distributions

Sampling Concepts

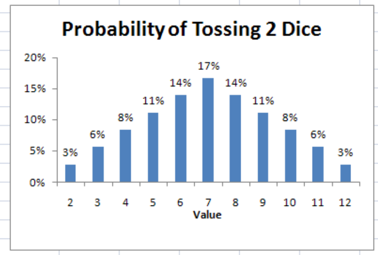
- Larger is better
- Larger is usually more costly
- Central limit theorem – sample means are normally distributed

Basis of Business Decisions

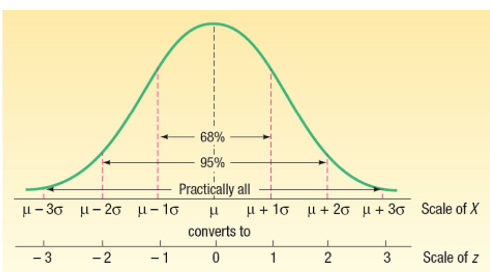


Tossing 2 Dice

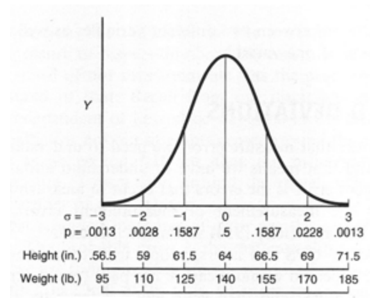
Value	Count	Relative
2	1	3%
3	2	6%
4	3	8%
5	4	11%
6	5	14%
7	6	17%
8	5	14%
9	4	11%
10	3	8%
11	2	6%
12	1	3%
	36	



Confidence Intervals – Continuous Normal Data



Another View of Normally Distributed Data



The standard normal curve.

A Marketing Example

Confidence Interval Estimates for the Mean

Use Z-distribution

- **If the population standard deviation is known or the sample is greater than 30.**

$$\bar{X} \pm z\frac{\sigma}{\sqrt{n}}$$

Use t-distribution

- **If the population standard deviation is unknown and the sample is less than 30.**

$$\bar{X} \pm t\frac{s}{\sqrt{n}}$$

+/- Error, E

Comparing the z and t Distributions when n is small

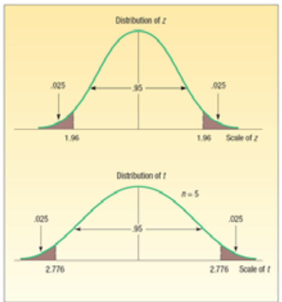


CHART 9–2 Values of z and t for the 95 Percent Level of Confidence

13

Calculate Sample Size

14

Sample Size Determination for a Variable

- To find the sample size for a variable:

$$n = \left(\frac{z \cdot s}{E} \right)^2$$

where:

E - the allowable error

z - the z - value corresponding to the selected level of confidence

s - the sample deviation (from pilot sample)

15

Problems with Sample Size

- Sigma is for population
- Need to know sigma for population

- Solutions (From Lind, p315)
- Use a comparable study
- Use a range-based approach (range/6)
- Conduct a pilot study

16

Calculating Sample Size

- Excel model

17

One more problem

- We may not know the population sigma
- Solution: Student-t distribution
- Problem: not explained by Lind (and others)
- Solution: Excel - Goal Seek

18

Why Dr. Shoemaker's Solution?

- p is for binomial data

$$n = p(1-p)\left(\frac{z}{E}\right)^2$$

Binomial Probability Distribution (discrete)

Characteristics:

- There are only two possible outcomes on a particular trial of an experiment.
- The outcomes are mutually exclusive,
- The random variable is the result of counts.
- Each trial is independent of any other trial
- Examples:
  - Yes or no
  - True or false
  - On or off
  - Correct or incorrect

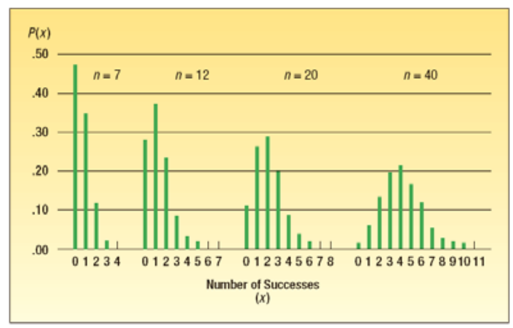
Binomial Probability Formula

BINOMIAL PROBABILITY FORMULA $P(x) = {}_nC_x\, \pi^x(1-\pi)^{n-x}$ [6–3]

where:

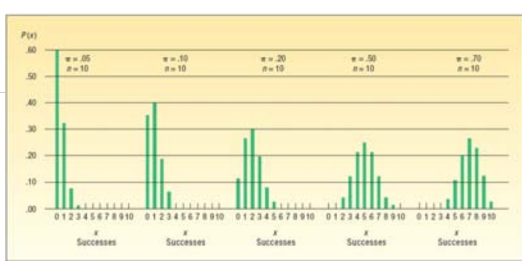
C denotes a combination.
n is the number of trials.
x is the random variable defined as the number of successes.
π is the probability of a success on each trial.

Binomial Distribution: n, x, π



Binomial – Shapes for Varying π (*n* constant)

CHART 6–2 Graphing the Binomial Probability Distribution for a π of .05, .10, .20, .50, and .70 and an n of 10



Some Interesting Characterists

- As n increases, binomial distribution starts looking normal
- As p gets close to .5, binomial distributions starts looking normal
- Given this, you can use the normal distribution to estimate the binomial

MEAN OF A BINOMIAL DISTRIBUTION $\mu = n\pi$ [6–4]

VARIANCE OF A BINOMIAL DISTRIBUTION $\sigma^2 = n\pi(1-\pi)$ [6–5]

Problems with this approach

- What is E?
- E is the allowable error, expressed as a percent (a value between 0 and 1)
- What is p?
- p is the probability of success. Most conservative estimate is .5

25

Equations for Sample Size

$$n = p(1-p)\left(\frac{z}{E}\right)^2 \qquad n = \left(\frac{z\sigma}{E}\right)^2$$

- How much will you pay for this cookie?
- Will it sell if it's price was $1.00?

26

27