Introductory Statistics Lectures

Binomial Distribution

Finding the probability of x successes in n trials.

Anthony Tanbakuchi
Department of Mathematics
Pima Community College

(Compile date: Tue May 19 14:49:27 2009)

Contents

1 Binomial Distribution

1.1 R Tip of the day

Calculating with R

If you use R to do a statistical calculation, use the following steps:

1. Determine which equation(s) you need to use.
2. Define variables in R with data for **all** the variables in your equation.
3. Type the equation into R.

Example 1. Given $x = \{4, 2, 7, 8\}$, find $q = \sqrt{\bar{x} - s}$

```
R: x = c(4, 2, 7, 8)
R: x.bar = mean(x)
R: s = sd(x)
R: q = sqrt(x.bar - s)
R: q
[1] 1.5799
```

1.2 Introduction

Question 1. For our class, 10 students wear corrective lenses and 8 do not. Find the probability of randomly selecting 4 students with replacement and 3 of the **4** wear corrective lenses.

1.3 Binomial distribution

DEFINITION 1.1

BINOMIAL DISTRIBUTION.

The probability of x successes in n trials with p probability of success is given by the binomial probability distribution:

$$\boxed{P(x \mid n, p) = {}_nC_x p^x q^{(n-x)}} \tag{1}$$

where ${}_nC_x$ is the number of ways you can choose x successes and $n-x$ failures in any order. The probability of failure is $q = 1 - p$.

Requirements:

1. Fixed number of trials n.
2. Independent trials: p remains constant for each trial.
   If sampling w/o replacement & $n/N \leq 0.05$ treat as independent.
3. Trial has 2 possible outcomes. (Y/N, T/F, blue/not blue)

USING THE DISTRIBUTION

Calculating binomial probability

A slightly different question.
For our class, 10 students wear corrective lenses and 8 do not. Find the probability of randomly selecting 10 students with replacement and 9 of the 10 do not wear corrective lenses.

Question 2. What does success represent?

Question 3. Determine what the parameters are: x, n, p

Two methods to compute probability

1. Use equation.
2. Use R function `dbinom()` .

Recall:

COMBINATIONS:
`choose(n,x)`
Finds ${}_nC_x$

R COMMAND

Question 4. Write out the equation to solve the problem.

Question 5. Use R as a calculator to solve the problem.

Question 6. Would it be unusual to observe 9 out of 10 students in our class not wearing corrective lenses?

R Command

BINOMIAL DISTRIBUTION:

`dbinom(x, n, p)`

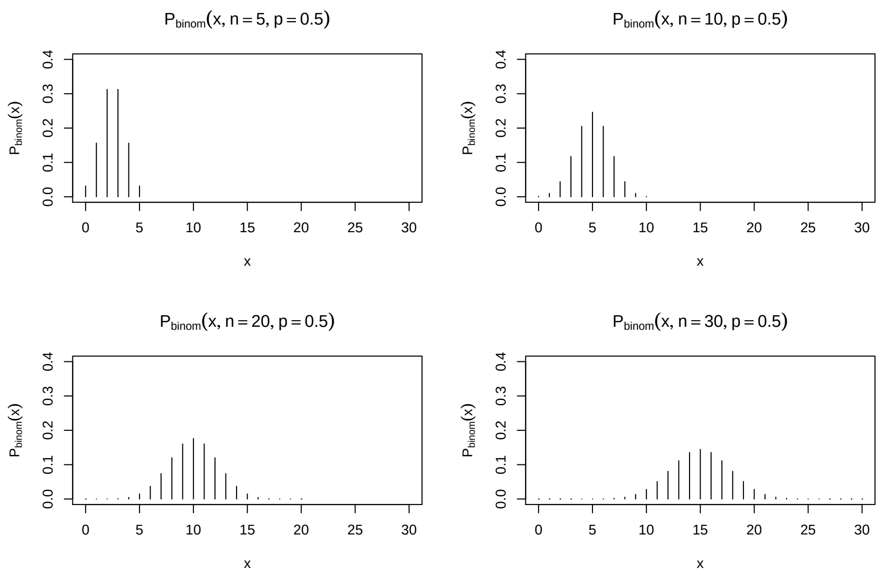
Finds the probability of `x` successes in `n` trials with `p` probability for individual success.

Question 7. Use the binomial distribution function in R to solve the problem.

DEPENDENCE ON N AND P

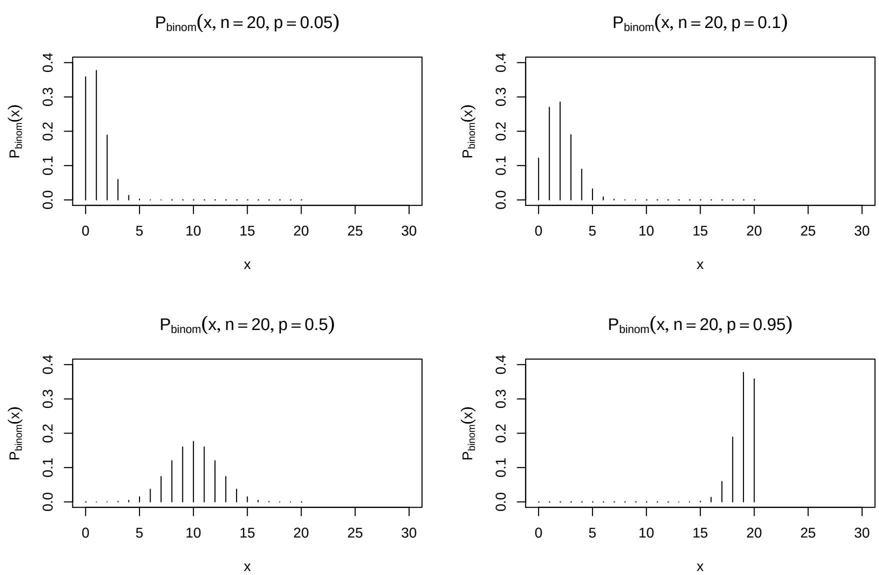
Binomial Distribution: dependence on n

Plot of binomial distribution with varying n, fixed $p = 0.5$.



Binomial Distribution: dependence on p

Plot of binomial distribution with fixed $n = 20$, varying p.



MEAN AND STANDARD DEVIATION

Mean of binomial distribution. Definition 1.2

$$\begin{aligned}\mu &= \sum_{i=1}^{k} x_i \cdot P(x_i) \\ &= \sum_{x=0}^{n} x \cdot {}_nC_x \cdot p^x q^{(n-x)} \\ &= np\end{aligned}$$

$$\boxed{\text{mean binomial } \mu = np} \tag{2}$$

Represents the mean number of successes x in n trials.

DEFINITION 1.3 STANDARD DEVIATION OF BINOMIAL DISTRIBUTION.

$$\begin{aligned}\sigma &= \sqrt{\sum_{i=1}^{k} (x_i - \mu)^2 \cdot P(x_i)} \\ &= \sqrt{\sum_{x=0}^{n} (x - \mu)^2 \cdot {}_nC_x \cdot p^x q^{(n-x)}} \\ &= \sqrt{npq}\end{aligned}$$

$$\boxed{\text{standard deviation binomial } \sigma = \sqrt{npq}} \tag{3}$$

Represents the standard deviation of number of successes x in n trials.

Again, for our class, 10 students wear corrective lenses and 8 do not. Answer the following questions if we randomly selecting 10 students with replacement and success represents not wearing corrective lenses..

Question 8. Find the mean μ of the binomial distribution.

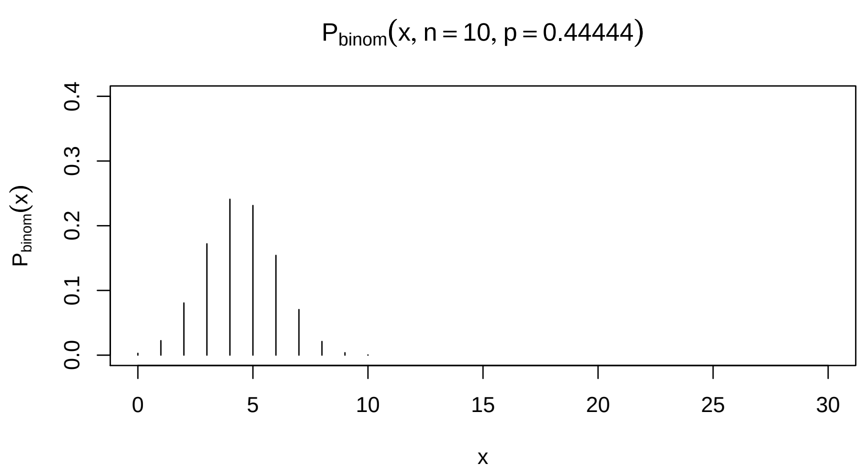
Question 9. What does the mean represent?

Question 10. Find the standard deviation σ of the binomial distribution.

Question 11. What does the standard deviation represent?

Question 12. What would be the usual number of successes we would expect using the Empirical Rule?

Visualizing μ and σ



MORE EXAMPLES

Determine if the binomial distribution applies to the following questions:

Question 13. Find the probability of 3 left handed students when randomly selecting 10 students from a class of 100 where 12 of them are left handed without replacement.

Question 14. Find the probability of 3 left handed students when randomly selecting 5 students from a class of 100 where 12 of them are left handed without replacement.

Question 15. 2% of Americans are ambidextrous. Find the probability of 3 ambidextrous students when randomly selecting 10 students from a class of 100 without replacement.

Example 2 (A worked out problem.). 2% of Americans are ambidextrous. Find the probability of 3 ambidextrous students when randomly selecting 10 students from a class of 100 without replacement.

(1) Binomial distribution applies (2) Find parameters.

```
R: x = 3
R: n = 10
R: p = 0.02
```

(3) Use the binomial distribution function

```
R: dbinom(x, n, p)
[1] 0.0008334
```

Example 3 (A worked out problem.). For our previous example, what is the probability of **3 or less** ambidextrous students in 10?

(1) Binomial distribution applies (2) Find parameters. Need to find the probability of x=0, x=1, x=2, x=3.

```
R: x = 0:3
R: x
[1] 0 1 2 3
R: P = dbinom(x, n, p)
R: P
[1] 0.8170728 0.1667496 0.0153137 0.0008334
```

Thus, the probability is the sum of of the probabilities for x=0 to 3:

```
R: sum(P)
[1] 0.99997
```

1.4 Related Distributions

GEOMETRIC DISTRIBUTION

GEOMETRIC DISTRIBUTION. DEFINITION 1.4

Describes the probability of observing the **first success** on the x^{th} trial for a set if independent trials with p probability of success for an individual trial.

$$P(x|p) = p \cdot q^{x-1} \tag{4}$$

Where on average the first success will be seen on trial:

$$\mu = \frac{1}{p} \tag{5}$$

Example 4. If we successively roll a die, what is the probability that the first time we observe a 2 is on the 10th trial?

```
R: x = 10
R: p = 1/6
R: P = p * (1 - p)^(x - 1)
R: signif(P, 3)
[1] 0.0323
```

The average number of trials until we first see a 2 would be: $\mu = 1/p = 6$.

HYPERGEOMETRIC DISTRIBUTION

HYPERGEOMETRIC DISTRIBUTION. DEFINITION 1.5

Like the binomial, but describes the probability of x successes in n trials **without replacement** from a population of size N with m total successes available.

$$P(x|n, N, m) = \frac{{}_mC_x \cdot {}_{N-m}C_{n-x}}{{}_NC_n} \tag{6}$$

$$= \frac{[\text{\# ways for } x \text{ successes}] \cdot [\text{\# ways for } n - x \text{ failures}]}{[\text{\# possible samples w/o replacement}]} \tag{7}$$

The average number of successes for a sample size n:

$$\mu = \frac{nm}{N} \tag{mean}$$

Example 5. 10 helmets are selected without replacement for destructive testing from a batch of 50 helmets containing 3 defective. What is the probability that 1 or more helmets is defective in the tested set?

[Success is a defective helmet.]

$$P(\text{1 or more}) = 1 - P(\text{none}) = 1 - P(x = 0|n = 10, N = 50, m = 3)$$

```
R: x = 0
R: n = 10
R: N = 50
R: m = 3
R: P = 1 - choose(m, x) * choose(N - m, n - x)/choose(N,
+     n)
R: signif(P, 3)
[1] 0.496
```

1.5 Summary

- Binomial Distribution: probability of x successes in n trials with p probability of success for an individual trial.
  1. Requirements:
     a) Fixed number of trials
     b) Independent trials (or $n/N \leq 0.05$)
     c) Two possible outcomes
  2. `P = dbinom(x, n, p)`
  3. Mean: $\mu = np$, "average number of successes x expected in n trials"
  4. Standard deviation: $\sigma = \sqrt{npq}$
- Geometric Distribution: Probability of first success on x^{th} trial.
- Hypergeometric Distribution: Probability of x successes in n trials **without replacement**.

1.6 Additional problems

Ten percent of American adults are left-handed. For a statistics class with 25 students answer the following:

Question 16. Find the probability that 3 out of 5 randomly selected students are left handed.

Question 17. Find the probability that 3 or fewer of 5 randomly selected students are left handed.

Question 18. Find the mean number of students you would expect to be left handed if 5 are randomly selected.

Question 19. Find the standard deviation of the number of left handed students when 5 are randomly selected.

Question 20. Would it be unusual to randomly select 5 left handed student?

Question 21. A new slot machine that lets you buy 20 plays for $20 will pay out $20,000 dollars if you get 3 lemons for the first time on the 10th play. If the probability you get three lemons when playing one turn on the machine is 0.0005, what is the probability that you win big on the 10th play?

Question 22. What are the average number of plays required until you win big on the slot machine?

Question 23. If you pick 5 students without replacement from a class, what is the probability that you select 5 males? (The class contains 12 females and 8 males)