# Introduction to Applied Categorical Data Analysis

## Edps/Psych/Soc 589

Carolyn J. Anderson

Department of Educational Psychology

ILLINOIS



Fall 2018

# Overview

- Data
- Sampling Models
  - Poisson Distribution
  - Binominal Distribution
- Inferences regarding proportions

# Data

- **Continuous/Numerical/Quantitative**: Measurement on these Variables are made on an interval or ratio scale. e.g.,
  - Income of a respondent to a survey
  - Height
  - Weight
  - Age
  - Reaction time
  - Blood pressure
  - Crop yield per acre
- **Categorical/Discrete/Qualitative:** Measures on categorical or discrete variables consist of assigning observations to one of a number of categories.

  The categories may be either
  - Unordered.
  - Ordered.

# Examples of Nominal Variables

There is no natural ordering of the categories; the ordering is arbitrary.

- Names of plants native to East Central Illinois
- Region of residence of survey respondent
- Telephone numbers
- Zip-codes
- Gender
- Occupation
- Race/ethincity
- Student's major field of study
- Received flu vaccination or not.
- Dead/alive
- High school program type
- Types of behaviors (e.g., aggressive, assertive, passive, etc.)

# Ordinal Variables

There is a natural ordering of the categories; the order is *not* arbitrary.

- Age group
- SES (high, middle, low)
- Response option selected on an item on an exam question (correct/incorrect).
- Degree of agreement with a statement on a questionaire.

# Examples of Ordinal Variables

- Some examples from the General Social Survey (GSS):

  "The public has little control over what politicians do in office."

  1. Agree
  2. Can't Choose
  3. Disagree
  4. (No Answer)

- "All employees should be required to retire at an age set by law"

  1. Agree strongly
  2. Agree
  3. Neither agree nor disagree
  4. Disagree
  5. Disagree strongly
  6. (No Answer)

# Nominal or Ordinal?

- Geroge Udny Yule:

  *"...all those who have died of small-pox are equally dead: no one of them is more or less dead than another, and the dead are are quite distinct from the survivors."*

- **Methods of Analysis.** Methods for nominal variables can be used to analysis nominal and/or ordinal variables, but methods for ordered variables are not appropriate for nominal data.
- **Context is important.** The categories of some variables may be ordered, partially ordered, or ordered with the appropriate ordering of the categories not known.

# Importance of Context

In one context, a variable may be ordinal but in another it may be nominal.

*"Did you get a vaccination for the flu?"*

- Yes
- No

**Nominal:** you either have or have not received a vaccine.
**Ordinal:** your perception of your susceptibility or the consequences of getting sick.

**Nominal:** different occupations.
**Ordinal:** consider the relationship between different occupations and SES.

# More Types of Data

**Counts:** Variables which represent a frequency of occurrence of an event.
Examples:

- Number of people who return a completed survey.
- Number of times the word "categorical" is used in today's lecture.
- Number of times specific behaviors are observed in a 10 minute taped session of two children playing together.
- Number of bacterial colonies on an agar plate.
- Number of accident on the corner of Green & 6th St during the 1st week of class.
- Number of correction answers on an exam.

# More Types of Data

**Proportion** or a "bounded count":

- The ratio of counts where the numerator equals the number of "successes" and the denominator equals that number of "trials".
- Includes *binary* data where the numerator is either 1 or 0 and the denominator is 1.
- Examples:
  - Number of students taking this class divided by the number of graduate students studying social sciences at UIUC.
  - Number of people who respond to a treatment out of those to whom it was administered.

# Another Classification of Variables

- **Response** or "dependent" variables. In this class, response variables are categorical, in particular, counts or proportions.
- **Explanatory** or "independent" variables. These may be continuous (numerical), ordinal, and/or nominal.

Whether a variable is a response or an explanatory variable depends on the context.

# Principles of Data Analysis

(from Wickens)

- A meaningful statistical analysis cannot be performed without non statistical information.
- The more non statistical information that is available, the stronger the conclusions that can be drawn.
- A confirmatory analysis yields conclusions that are both stronger and more precise than those of an exploratory analysis.

  Confirmatory analysis — question driven.
  Exploratory analysis — data driven.

- A precisely formulated question gives rise to a specific statistical analysis; whereas weak ones give little direction.

# Sampling Models for Categorical Data

Random Mechanisms

Two important ones that we'll start with are

- Poisson Sampling
- Binomial Sampling

# Poisson Sampling

The Poisson distribution...

- Is used for counts of events that occur randomly over time (or space).
- Is often useful when the probability of event on any particular trial is very small while the number of trials is very large.
- Arises naturally from counting random events during a fixed period of time.

**Requirements:**

1. Events must be independent.
2. Time period (or space) must be fixed.

# Examples of Poisson Sampling

- The number of suicides in the US in a year.
- Number of e-mails received between 8am and 9am Monday.
- Number of flaws in 100 feet of wire.
- Bortkiewicz (1898) (or Bortkewicz or Bortkewitsch or Bortkewitch ?).

  The event: yearly total of men in the Prussian army corps who were kicked by horses and died as a results of their injuries. 20 year periods and collected on 14 different corps.

| Number Killed | Observed Frequency |
|---|---|
| 0 | 144 |
| 1 | 91 |
| 2 | 32 |
| 3 | 11 |
| 4 | 2 |
| 5+ | 0 |
| total | 280 |

# 1994 World Cup Soccer

(from the internet).

Event = frequency of various number of goals scored by a team during the 1st round of play (out of 35 matches).

| Number of Goals | Observed Frequency |
|---|---|
| 0 | 20 |
| 1 | 29 |
| 2 | 16 |
| 3 | 3 |
| 4 | 1 |
| 5 | 0 |
| 6 | 1 |
| 7+ | 0 |
| Total | 70 |

# Poisson Process

More formal requirements.. . . 3 rules:

- The number of events (or "changes") in non-overlapping intervals are independent.
- The probability of exactly 1 event is proportional to the length of the interval. More technically, the probability of exactly 1 event occurring in a "sufficiently" short interval of length $\Delta t$ is approximately $\Delta t \pi$.
- The probability of 2 or more events in a sufficiently short interval is essentially zero.

# Example of Poisson

The number of people who arrive at the Illini Union bookstore during a 5 minute period during the 1st week of classes to buy books.

Why might Poisson distribution describe this?

Divide time (the 5 minutes) into tiny-tiny intervals. Then the probability that a person arrives in any particular time interval is very small, but in 5 minutes the number of people who arrive can be very large.

# Poisson Distribution

The distribution function is

$$P(y) = \frac{e^{-\mu}\mu^{y}}{y!} \qquad\qquad y = 0, 1, 2, \ldots$$

where

- $y!$ is "$y$ factorial"

$$y! = 1 \times 2 \times 3 \times \ldots \times y.$$

and

$$0! = 1$$

- $\mu$ is the *parameter* of the distribution. Once you know $\mu$, you know everything there is to know about the distribution.
- $e^{-\mu}$ is the exponential function evaluated at $-\mu$.
  Note:
  - $e = 2.718\ldots$
  - If $e^{a} = b$, then $\ln(b) = a$ where "ln" is the *natural* logarithm.
  - In this class, whenever your see "log" it refers to the natural log or "ln".

# World Cup Soccer Revisited

Suppose $\mu = 1.143$, so $P(y) = \frac{e^{-1.143}(1.143)^y}{y!}$

$$
\begin{aligned}
P(0) &= \frac{e^{-1.143}(1.143)^0}{0!} = \frac{(.139)(1.000)}{1} = .319 \\
P(1) &= \frac{e^{-1.143}(1.143)^1}{1!} = \frac{(.139)(1.143)}{1} = .364 \\
P(2) &= \frac{e^{-1.143}(1.143)^2}{2!} = \frac{(.139)(1.306)}{2} = .208 \\
P(3) &= \frac{e^{-1.143}(1.143)^3}{3!} = \frac{(.139)(1.493)}{6} = .079 \\
P(4) &= \frac{e^{-1.143}(1.143)^4}{4!} = \frac{(.139)(1.707)}{24} = .023 \\
P(5) &= \frac{e^{-1.143}(1.143)^5}{5!} = \frac{(.139)(1.951)}{120} = .005 \\
P(6) &= \frac{e^{-1.143}(1.143)^6}{6!} = \frac{(.139)(2.223)}{720} = .001 \\
P(7) &= \text{a small number}
\end{aligned}
$$

# World Cup Soccer Expected Frequencies

We can compute expected frequencies:

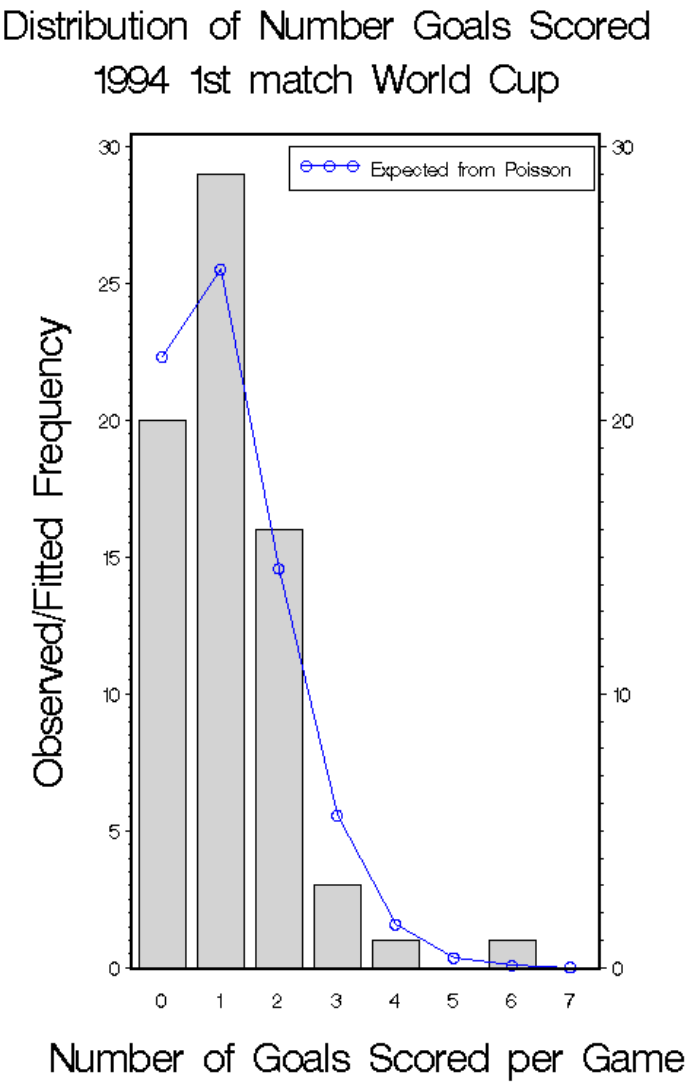
$$\text{Expected frequency that } (Y = y) = 70P(Y = y)$$

$$\text{e.g., Expected frequency that } (Y = 0) = 70(.319) = 22.32$$

| Number of Goals | Observed Frequency | Probability $P(y)$ | Expected Frequency |
|---|---|---|---|
| 0 | 20 | .319 | 22.324 |
| 1 | 29 | .364 | 25.513 |
| 2 | 16 | .208 | 14.579 |
| 3 | 3 | .079 | 5.554 |
| 4 | 1 | .023 | 1.587 |
| 5 | 0 | .005 | 0.363 |
| 6 | 1 | .001 | 0.069 |
| 7+ | 0 | .000 | 0.011 |
| Totals | 70 | 1.00 | 70.00 |

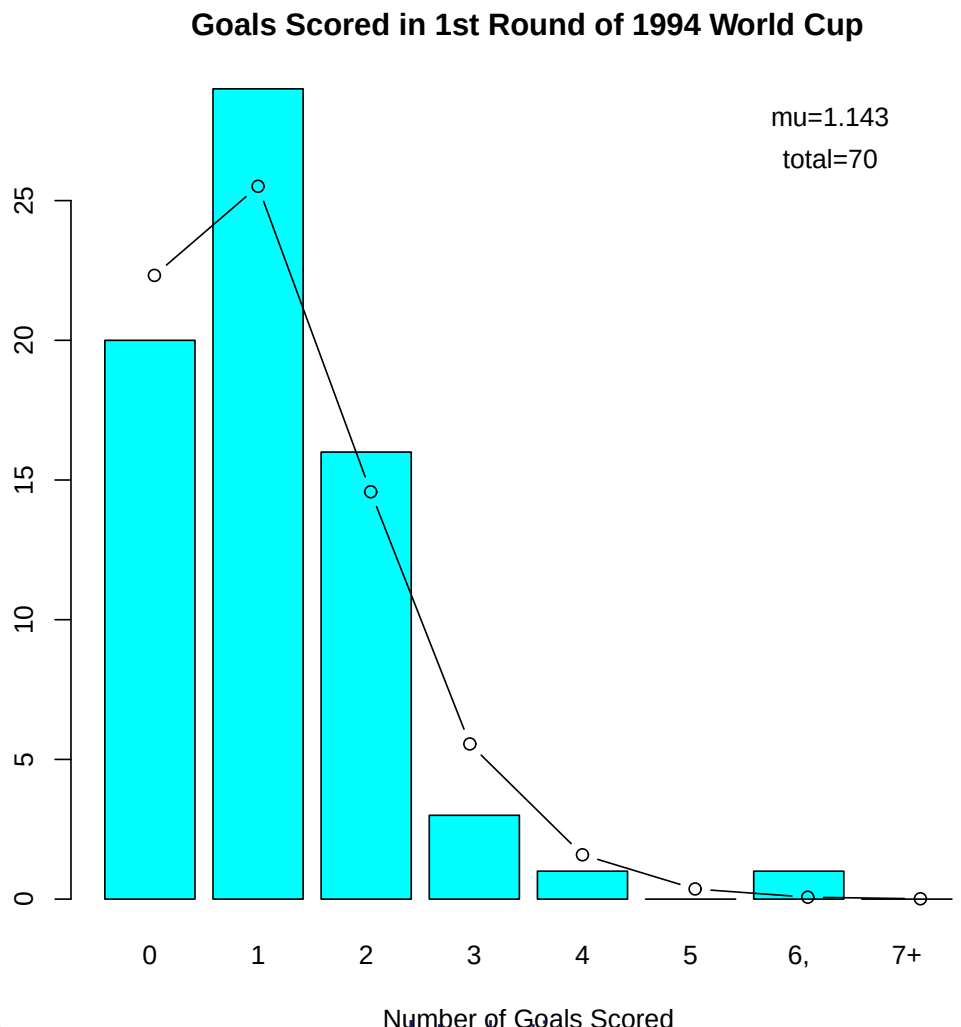
# 1994 World Cup Soccer Figure

1994 data & fitted Poisson:



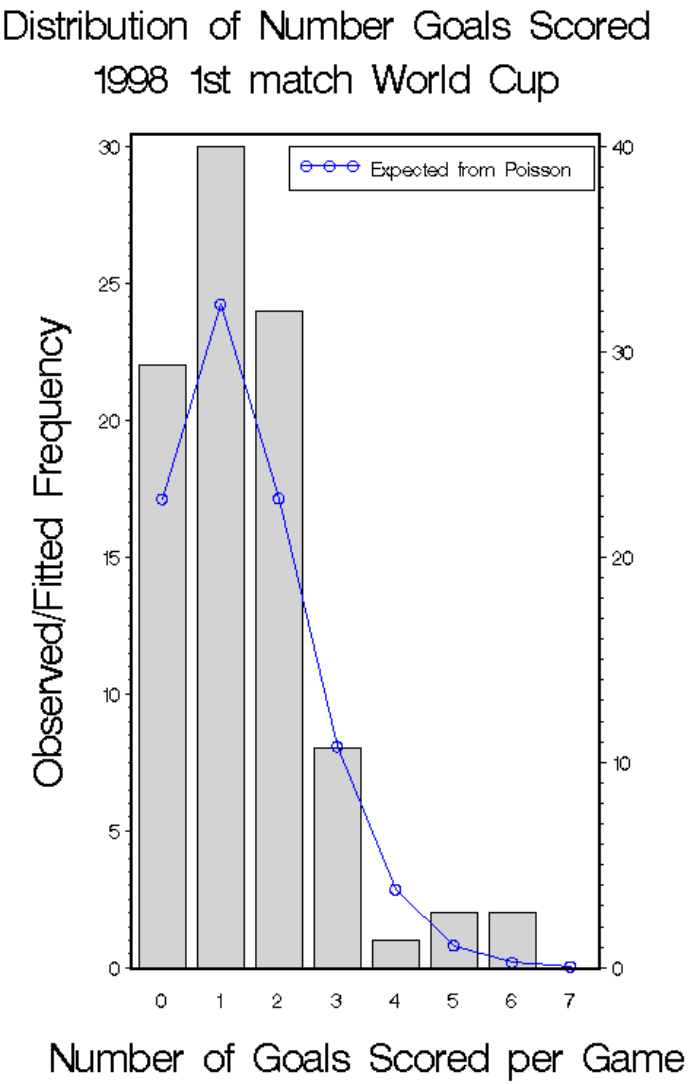
# R Figure

1994 data & fitted Poisson:



# 1998 World Cup Soccer Figure

1998 data & fitted Poisson:



# Sample Statistics for 94 World Cup Soccer

The mean number of goals (per team) equals

$$\begin{aligned}\bar{y} &= \frac{1}{70}\sum_{i=1}^{70} y_i \\ &= \frac{1}{70}(\overbrace{0+0+\ldots+0}^{20}+\overbrace{1+\ldots+1}^{22}+\ldots+6) \\ &= 1.143\end{aligned}$$

Note: $y_i$ is the number of goals that team $i$ scored.

The variance equals

$$\text{var}(y) = \frac{1}{70}\sum_{i=1}^{70}(y_i-\bar{y})^2 = 1.168$$

The mean and variance of $y$ are about equal.

# Important Property of the Poisson Distribution

The Mean = The Variance

$$E(Y) = \sigma^2$$

$$\mu = \sigma^2$$

In practice, the variance is usually larger than the mean. This is called *overdispersion*

In the soccer example, we (implicitly) assumed

- The mean number of goals scored by a team is the same for all teams; the "*Homogeneity*" assumption.
- The number of goals scored by a team is independent of the number of goals scored by other teams; the "*Independence*" assumption.

# The Binomial Distribution

We now assume that the number of trials is fixed and we count the number of "successes" or events that occur.

Preliminaries: *Bernoulli random variables*

- $X$ is a random variable where $X = 1$ or 0
- The probability that $X = 1$ is $\pi$
- The probability that $X = 0$ is $(1 - \pi)$

Such variables are called *Bernoulli random variables*.

# Bernoulli Random Variable

The mean of a Bernoulli random variable is

$$\mu_x = E(X) = 1\pi + 0(1-\pi) = \pi$$

The variance of $X$ is

$$\begin{aligned} \text{var}(X) = \sigma_X^2 &= E[(X-\mu_X)^2] \\ &= (1-\pi)^2\pi + (0-\pi)^2(1-\pi) \\ &= \pi(1-\pi) \end{aligned}$$

# Example of Bernoulli Random Variable

Suppose that a coin is

- "not fair" or is "loaded"
- The probability that it lands on heads equals .40 and the probability that it lands on tails equals .60.
- If this coin is flipped many, many, many times, then we would expect that it would land on heads 40% of the time and tails 60% of the time.
- We define our Bernoulli random variable as

$$X = \begin{cases} 1 & \text{if Heads} \\ 0 & \text{if Tails} \end{cases}$$

**Note:** where $P(X = 1) = .40$ and $P(X = 0) = .60$. Once you know $\pi$, you know the mean and variance of the distribution of $X$.

# Binomial Distribution

A binomial random variable is the sum of $N$ independent Bernoulli random variables. We will let $Y$ represent a binomial random variable and by definition

$$Y = \sum_{i=1}^{N} X_i$$

The mean of a Binomial random variable is

$$\begin{aligned}
\mu_y = E(Y) &= E(\sum_{i=1}^{N} X_i) \\
&= E(X_1) + E(X_2) + \ldots + E(X_N) \\
&= \overbrace{\mu_x + \mu_x + \ldots + \mu_x}^{N} \\
&= \overbrace{\pi + \pi + \ldots + \pi}^{N} \\
&= N\pi
\end{aligned}$$

# Variance of Binomial Random Variable

. . . and the variance of a Binomial random variable is

$$\begin{aligned}
\text{var}(Y) = \sigma_y^2 &= \text{var}(X_1 + X_2 + \ldots + X_N) \\
&= \overbrace{\text{var}(X) + \text{var}(X) + \ldots + \text{var}(X)}^{N} \\
&= \overbrace{\pi(1-\pi) + \pi(1-\pi) + \ldots + \pi(1-\pi)}^{N} \\
&= N\pi(1-\pi)
\end{aligned}$$

**Note:** Once you know $\pi$ and $N$, you know the mean and variance of the Binomial distribution.

# Binomial Distribution Function

- Toss the unfair coin with $\pi = .40$ coin $N = 3$ times.
- $Y$ = number of heads.
- The tosses are independent of each other.

By multiplication rule & addition rule from probability theory

| Possible Outcomes $X_1 + X_2 + X_3 = Y$ | Probability of a Sequence $P(X_1, X_2, X_3)$ | Prob(Y) $P(Y)$ |
|---|---|---|
| $1+1+1=3$ | $(.4)(.4)(.4) = (.4)^3(.6)^0 = .064$ | .064 |
| $1+1+0=2$ | $(.4)(.4)(.6) = (.4)^2(.6)^1 = .096$ | |
| $1+0+1=2$ | $(.4)(.6)(.4) = (.4)^2(.6)^1 = .096$ | $3(.096) = .288$ |
| $0+1+1=2$ | $(.6)(.4)(.4) = (.4)^2(.6)^1 = .096$ | |
| $1+0+0=1$ | $(.4)(.6)(.6) = (.4)^1(.6)^2 = .144$ | |
| $0+1+0=1$ | $(.6)(.4)(.6) = (.4)^1(.6)^2 = .144$ | $3(.144) = .432$ |
| $0+0+1=1$ | $(.6)(.6)(.4) = (.4)^1(.6)^2 = .144$ | |
| $0+0+0=0$ | $(.6)(.6)(.6) = (.4)^0(.6)^3 = .216$ | .216 |
| | 1.000 | 1.000 |

# Binomial Distribution Function

The formula for the probability of a Binomial random variable is

$$P(Y=a) = \left( \begin{array}{c} \text{the number of ways that} \\ Y=a \text{ out of } N \text{ trials} \end{array} \right) P(X=1)^a P(X=0)^{(N-a)}$$

$$= \left( \begin{array}{c} N \\ a \end{array} \right) \pi^a (1-\pi)^{N-a}$$

where

$$\left( \begin{array}{c} N \\ a \end{array} \right) = \frac{N!}{a!(N-a)!} = \frac{N(N-1)(N-2)\ldots 1}{a(a-1)\ldots 1((N-a)(N-a-1)\ldots 1)}$$

which is called the "binomial coefficient."

For example, the number of ways that you can get $Y=2$ out of 3 tosses is

$$\left( \begin{array}{c} 3 \\ 2 \end{array} \right) = \frac{3(2)(1)}{2(1)(1)} = 3$$

# Statistical Inferences Regarding Probability, $\pi$

Example: Data from Sommers (2000) *Chance*.

The data are the number of times that the taller candidate (from two major parties) of won the US presidential election. We'll just consider elections from 1932 to 1992 (Clinton vs Bush).

| Winner was | Taller | Shorter | Total |
|---|---|---|---|
| | $y$ | $N-y$ | $N$ |
| | 13 | 2 | 15 |

The observed proportion of times that taller candidate won:

$$p = 13/15 = .8667$$

Is this significantly different from chance?

$$H_o : \pi = .5 \qquad \text{versus} \qquad H_a : \pi \neq .5$$

. . . but first. . .

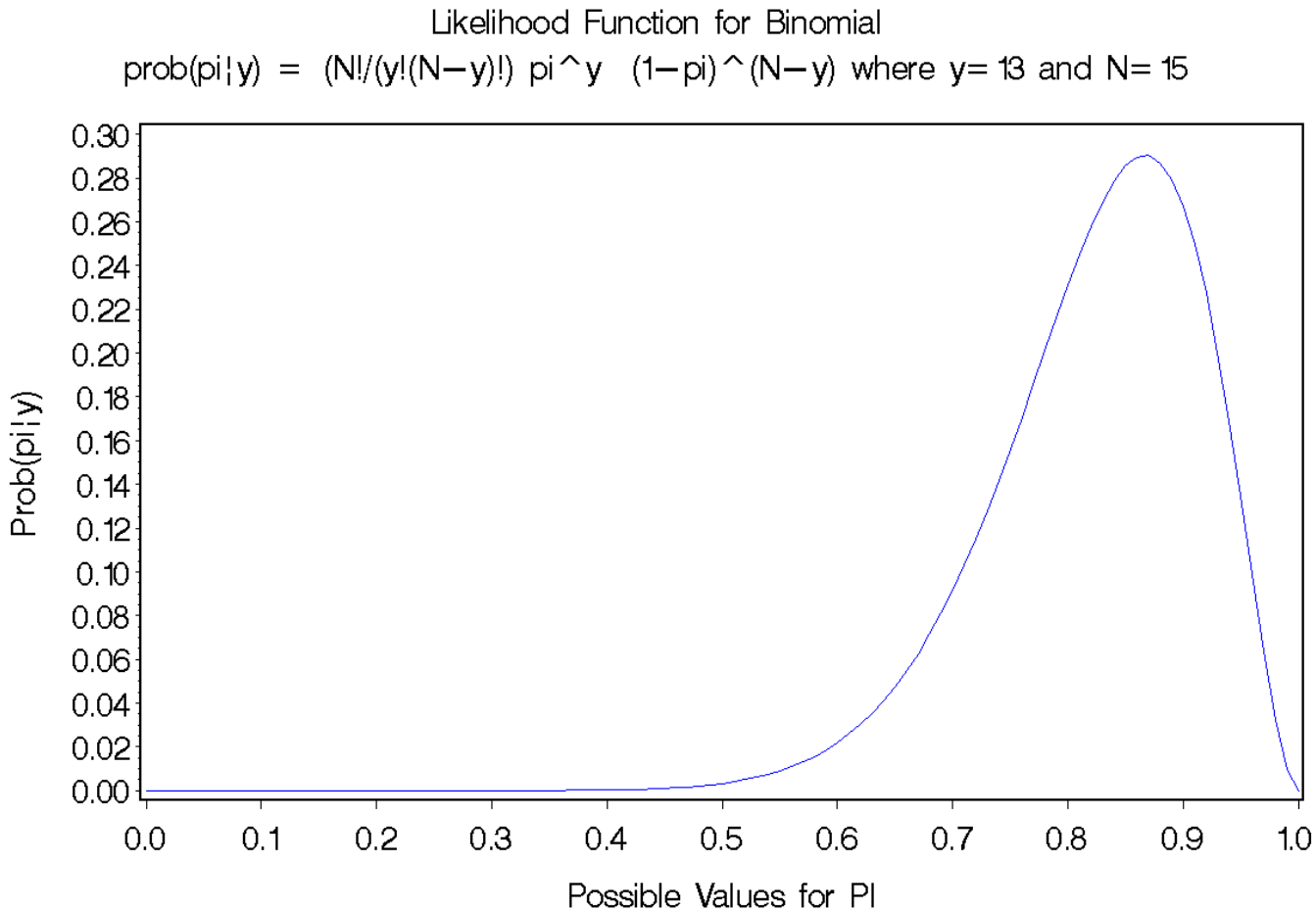
# Maximum Likelihood Estimation

- The Likelihood Function for Binomial

$$P(\pi|y) = \frac{N!}{y!(N-y)!}\pi^y(1-\pi)^{(N-y)}$$

  where $y$ is known and $\pi$ is unknown.
- $P(\pi|y)$ is now how likely the population proportion equals $\pi$ given the data $y$.
- The value of $\pi$ that is the *most likely* given the data is the "maximum likelihood estimate" of $\pi$.
- Denote MLE's by "^", for example $\hat{\pi}$.
- For binomial: $\hat{\pi} = y/N$ (and $\hat{\mu} = \hat{\pi}N$).
- . . . and for Poisson: $\hat{\mu} = \hat{\sigma}^2 = (1/I)\sum_{i=1}^{I} y_i$.

# MLE Function for Presidential Data



# Presidential Election Data

Hypothesis test $\rightarrow H_o : \pi = .5$ versus $H_a : \pi \neq .5$

Using the Binomial distribution,

$$\hat{\pi} = p = y/N = 13/15 = .8667$$

Need the probability of $y = 13$ (or $p = .8667$) or something more extreme assuming that the null is true — the $p$–value.

$$\begin{aligned}
P(15) &= \frac{15!}{15!0!}(.5)^{15}(1-.5)^{0} = .000031 \\
P(14) &= \frac{15!}{14!1!}(.5)^{14}(1-.5)^{1} = .000457 \\
P(13) &= \frac{15!}{13!2!}(.5)^{13}(1-.5)^{2} = .003204 \\
P(2) &= \frac{15!}{2!13!}(.5)^{2}(1-.5)^{13} = .003204 \\
P(1) &= \frac{15!}{1!14!}(.5)^{1}(1-.5)^{14} = .000457 \\
P(0) &= \frac{15!}{0!15!}(.5)^{0}(1-.5)^{15} = .000031
\end{aligned}$$

$p$–value $= .0074$

# The Sampling Distribution of Proportions

The observed proportion, $p$, equals $\hat{\pi}$:

$$p = \hat{\pi} = \frac{Y}{N}$$

Proportions are sample means of the Bernoulli random variables,

$$p = \frac{Y}{N} = \frac{1}{N}\sum_{i=1}^{N} X_i = \bar{X}$$

The mean of the sampling distribution of $p$:

$$\begin{aligned}
\mu_p = E(p) &= E(\frac{1}{N}\sum_{i=1}^{N} X_i) \\
&= \frac{1}{N}(E(X_1) + E(X_2) + \ldots + E(X_N)) \\
&= \frac{1}{N}(\overbrace{\pi + \pi + \ldots + \pi}^{N}) \\
&= \frac{1}{N}(\pi N) = \pi = \mu_x \longrightarrow \text{unbiased}
\end{aligned}$$

# The Sampling Distribution of Proportions

The variance of $p$ is

$$\begin{aligned} \text{var}(p) = \sigma_p^2 &= \text{var}(\frac{Y}{N}) \\ &= \frac{1}{N^2}\text{var}(Y) \\ &= \frac{N\pi(1-\pi)}{N^2} = \frac{\pi(1-\pi)}{N} = \frac{\text{var}(X)}{N} \end{aligned}$$

The shape of the sampling distribution?

By the central limit theorem, if $N$ is "large enough", then $p$ is approximately distributed as a normal random variable with mean $\pi$ and variance $\pi(1-\pi)/N$;

$$p \approx \mathcal{N}(\pi, \pi(1-\pi)/N)$$

When the "parent" distribution is Binomial, "large enough" usually means that $N\pi \geq 5$ and $N(1-\pi) \geq 5$.

# Large Sample Tests of Hypotheses Regarding $\pi$

We can use z and the standard normal distribution:

$$z = \frac{(p - \pi_o)}{\sqrt{\pi_o(1 - \pi_o)/N}}$$

- $p$ is the observed proportion of occurrences of the event.
- $\pi_o$ is the null hypothesized probability.
- $\sqrt{\pi_o(1 - \pi_o)/N}$ is the standard error of the sampling distribution of $p$ (under $H_o$).

In our Presidential Election example

$$z = \frac{(.8667 - .5)}{\sqrt{.5(1 - .5)/15}} = 2.84$$

The p-value for $z = 2.84$ is .005.
The exact $p$–value equals .007. If $N$ was larger, then the normal distribution would be a better approximation of the sampling distribution of $p$ and the $p$–values would be closer in value.

$$15(13/15) = 13 > 5, \text{ but } 15(2/15) = 2 < 5.$$

# Method I: Confidence Interval for $\pi$

The usual method for $(1-\alpha)\times 100\%$ CI,

$$p \pm z_{\alpha/2}\sqrt{p(1-p)/N}$$

A 95% Presidential Data:

$$p = .8667 \pm 1.96\sqrt{.8667(1-.8667)/15} \longrightarrow (.69, 1.04)$$

Need a better method for $(1-\alpha)\times 100\%$ CI

# Method II: Confidence Interval for $\pi$

Consider all $\pi_o$ for which you would *not* reject the null hypothesis,

$$-z_{\alpha/2} = \frac{p - \pi_o}{\sqrt{\pi_o(1-\pi_o)/N}} \quad \text{and} \quad \frac{p - \pi_o}{\sqrt{\pi_o(1-\pi_o)/N}} = z_{\alpha/2}$$

The solution for $\pi_o$:

$$\begin{aligned} a &= 1 + z^2_{\alpha/2}/N \\ b &= -2p - z^2_{\alpha/2}/N \\ c &= p^2 \end{aligned}$$

and

$$\pi_o = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

95% CI for the Presidential Election Data: $(.621, .963)$.

# Another Example: 1994 respondents to the GSS

"Please tell me whether or not you think it should be possible for a pregnant woman to obtain a legal abortion if the family has a very low income and cannot afford any more children".

| Response | Number |
|---|---|
| No | 954 |
| Yes $y =$ | 971 |
| Total $N =$ | 1925 |

$$H_o : \pi = .5 \text{ versus } H_a : \pi \neq .5$$

$$z = \frac{971/1925 - .5}{\sqrt{.5(1-.5)/1925}} = .387$$

$p$–value $= .698$.

A 95% CI for $\pi$ equals $(.482, .527)$.

# One more example: 2000 Presidential Election

The 2000 US presidential election came down to votes in Florida. The official results from the Florida Department of State, Division of Elections for the two top candidates as of Sunday November 28, 2000 are

| | |
|---|---|
| George W. Bush | 2,912,790 |
| Al Gore | 2,912,253 |
| Total | 5,825,043 |

This gave George W. Bush a 537 vote lead.

Observed proportion for Bush:

$$p = \frac{2,912,790}{5,825,043} = .500046094$$

$H_o : \pi = .5$; that is, the election was a tie versus $H_o : \pi \neq .5$.

$$z = \frac{.500046094 - .5}{\sqrt{\frac{.5(.5)}{5,825,043}}} = \frac{.000046094}{.00207166} = .0222$$

p-value = .98.

# 99% CI for $\pi$

99% CI for the probability that George W. won:

$$.500046094 \pm 2.576(.000207166) \Longrightarrow (.4995, .5006)$$

Given that the total number of votes equaled 5,825,043, how many votes would a candidate have needed such that the probability that the candidate won equaled .99?

i.e., where Power $= .99$?

# Margin for Victory

Test statistic (note: ${}_{.995}z = 2.576$)

$$2.576 = \frac{p - .50}{.00207166}$$

After re-arranging terms,

$$p = 2.576(.00207166) + .5 = .500533662$$

$$\begin{aligned} \text{So, the number of votes for winner} &= Np \\ &= 5,825,043(.500533662) \\ &= 2,915,630.104 \end{aligned}$$

$$\begin{aligned} \text{number of votes for loser} &= 5,825,043 - 2,915,630.104 \\ &= 2,909,412.896 \end{aligned}$$

The margin of victory needed for power = .99 would be

$$2,915,630.104 - 2,909,412.896 = 6,217.208$$

# SAS & Tests for Proportions

Out of 15 elections, 13 of taller candidates won, which gave us $p = 13/15 = .8667$.

To test $H_o : \pi = .5$ in SAS:

```
DATA height;
   INPUT height $ count;
   DATALINES;
    taller 13
    shorter 2
RUN;
PROC FREQ DATA=height ORDER=data;
   WEIGHT count;
   TABLES height / binomial  (p=.5);
RUN;
```

# SAS OUTPUT

Presidential Election data: Test for proportion

The FREQ Procedure

| height | Frequency | Percent | Cumul. Freq | Cumul. Percent |
|---|---|---|---|---|
| taller | 13 | 86.67 | 13 | 86.67 |
| shorter | 2 | 13.33 | 15 | 100.00 |

# SAS OUTPUT

```
     Binomial Proportion
     for height = taller
Proportion                 0.8667
ASE                        0.0878
95% Lower Conf Limit       0.6946
95% Upper Conf Limit       1.0000

Exact Conf Limits
95% Lower Conf Limit       0.5954
95% Upper Conf Limit       0.9834
```

# SAS OUTPUT

```
Test of H0:  Proportion = 0.5

ASE under H0                0.1291
Z                           2.8402
One-sided Pr > Z            0.0023
Two-sided Pr > |Z|          0.0045

Sample Size = 15
```