

UCLA STAT 13
**Introduction to Statistical Methods for
the Life and Health Sciences**

- **Instructor:** Ivo Dinov,
Asst. Prof. In Statistics and Neurology
- **Teaching Assistants:**
Ming Zheng, Sovia Lau, Tianwei Yu
UCLA Statistics

University of California, Los Angeles, Fall 2003
http://www.stat.ucla.edu/~dinov/courses_students.html

STAT 13, UCLA, Ivo Dinov Slide 1

Chapter 5: Discrete Random Variables

- Random variables
- Probability functions
- The Binomial distribution
- Expected values

STAT 13, UCLA, Ivo Dinov Slide 2

Definitions

- An **experiment** is a naturally occurring **phenomenon**, a scientific **study**, a sampling **trial** or a **test.**, in which an object (unit/subject) is selected at random (and/or treated at random) to *observe/measure* different outcome characteristics of the process the experiment studies.
- A **random variable** is a type of measurement taken on the outcome of a random experiment.

Slide 3 STAT 13, UCLA, Ivo Dinov

Definitions

- The **probability function** for a discrete random variable X gives $P(X = x)$ [denoted $pr(x)$ or $P(x)$] for every value x that the R.V. X can take
- E.g., number of heads when a coin is tossed twice

x	0	1	2
$pr(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Slide 4 STAT 13, UCLA, Ivo Dinov

Stopping at one of each or 3 children

Sample Space – complete/unique description of the possible outcomes from this experiment.

Outcome	GGG	GGB	GB	BG	BBG	BBB
Probability	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$

- For R.V. $X =$ number of girls, we have

X	0	1	2	3
$pr(x)$	$\frac{1}{8}$	$\frac{5}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

Slide 5 STAT 13, UCLA, Ivo Dinov

Plotting the probability function

X	0	1	2	3
$pr(x)$	$\frac{1}{8}$	$\frac{5}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

Slide 6 STAT 13, UCLA, Ivo Dinov

Tossing a biased coin twice

- For each toss, $P(\text{Head}) = p \rightarrow P(\text{Tail}) = P(\text{comp(H)}) = 1-p$
- Outcomes: HH, HT, TH, TT
- Probabilities: $p, p, p(1-p), (1-p)p, (1-p)(1-p)$
- Count X , the number of heads in 2 tosses

X	0	1	2
$\text{pr}(x)$	$(1-p)^2$	$2p(1-p)$	p^2

Slide 7 STAT 13, UCLA, Joe Dineen

Hospital stays

Days stayed	x	4	5	6	7	8	9	10	Total
Frequency		10	30	113	79	21	8	2	263
Proportion $\text{pr}(X=x)$		0.038	0.114	0.430	0.300	0.080	0.030	0.008	1.000
Cumulative Proportion $\text{pr}(X \leq x)$		0.038	0.152	0.582	0.882	0.962	0.992	1.000	

From Chance Encounters by C.J. Wild and G.A.F. Seber, © John Wiley & Sons, 2000.

Slide 8 STAT 13, UCLA, Joe Dineen

Calculating Interval probabilities from cumulative probabilities

x -values : 1 2 3 4 5 6 7 8 9 10 11 12
 To get 4 to 8, $P(3 < X \leq 9)$
 start with everything up to 8, $P(X \leq 9)$
 and remove from 3 down, $P(X \leq 3)$

Figure 5.2.2 Interval probabilities from cumulative probabilities.^a
[This Figure represents an arbitrary distribution, not the hospital distribution.]

From Chance Encounters by C.J. Wild and G.A.F. Seber, © John Wiley & Sons. **How to find the upper-tail?**

Slide 9 STAT 13, UCLA, Joe Dineen

Review

- What is a **random variable**? What is a discrete random variable? (type of measurement taken on the outcome of random experiment)
- What general principle is used for finding $P(X=x)$? (Adding the probabilities of all outcomes of the experiment where we have measured the RV, $X=x$)
- What two general properties must be satisfied by the probabilities making up a probability function? ($P(x) \geq 0$; $\sum P(x) = 1$)
- What are the two names given to probabilities of the form $P(X \leq x)$? (cumulative & lower/left-tail)

Slide 10 STAT 13, UCLA, Joe Dineen

Review

- How do we find an **upper/right-tail** probability from a **cumulative** probability? [$P(X > x) = 1 - P(X \leq x)$]
- When we use $P(X \leq 12) - P(X \leq 5)$ to calculate the probability that X falls within an **interval** of values, what numbers are included in the interval? ((6:12))

Slide 11 STAT 13, UCLA, Joe Dineen

The two-color urn model

N balls in an urn, of which there are

- M black balls
- $N - M$ white balls

Sample n balls and count $X = \#$ black balls in sample

We will compute the probability distribution of the R.V. X

Slide 12 STAT 13, UCLA, Joe Dineen

The biased-coin tossing model

toss 1 toss 2 ... toss n
 $\text{pr}(H) = p$ $\text{pr}(H) = p$ $\text{pr}(H) = p$

Perform n tosses and count $X = \#$ heads

We also want to compute the probability distribution of this R.V. X !
Are the two-color urn and the biased-coin models related? How do we present the models in mathematical terms?

Slide 13 STAT 13, UCLA, Jon Dineen

The answer is: **Binomial distribution**

- The distribution of the number of heads in n tosses of a biased coin is called the **Binomial distribution**.

Slide 14 STAT 13, UCLA, Jon Dineen

Binomial(N, p) – the probability distribution of the number of Heads in an N -toss coin experiment, where the probability for Head occurring in each trial is p .
E.g., Binomial(6, 0.7)

	x	0	1	2	3	4	5	6
Individual	$\text{pr}(X=x)$	0.001	0.010	0.060	0.185	0.324	0.303	0.118
Cumulative	$\text{pr}(X \leq x)$	0.001	0.011	0.070	0.256	0.580	0.882	1.000

For example $P(X=0) = P(\text{all 6 tosses are Tails}) = (1 - 0.7)^6 = 0.3^6 = 0.001$

Slide 15 STAT 13, UCLA, Jon Dineen

Binary random process

The *biased-coin tossing model* is a physical model for situations which can be characterized as a series of trials where:

- each trial has only **two outcomes**: *success* or *failure*;
- $p = P(\text{success})$ is the same for every trial; and
- trials are **independent**.

- The distribution of $X =$ number of successes (heads) in N such trials is

Binomial(N, p)

Slide 16 STAT 13, UCLA, Jon Dineen

Sampling from a finite population – Binomial Approximation

If we take a sample of size n

- from a much larger population (of size N)
- in which a proportion p have a characteristic of interest, then the distribution of X , **the number in the sample with that characteristic**,
- is *approximately* Binomial(n, p).
 - (Operating Rule: Approximation is adequate if $n/N < 0.1$.)
- Example, polling the US population to see what proportion is/has-been married.

Slide 19 STAT 13, UCLA, Jon Dineen

Odds and ends ...

- For what types of situation is the urn-sampling model useful? For modeling binary random processes. When sampling **with replacement**, Binomial distribution is **exact**, where as, in sampling **without replacement** Binomial distribution is an **approximation**.
- For what types of situation is the biased-coin sampling model useful? Defective parts. Approval poll of cloning for medicinal purposes. Number of Boys in 151 presidential children (90).
- Give the three essential conditions for its applicability. (**two outcomes**; **same p** for every trial; **independence**)

Slide 20 STAT 13, UCLA, Jon Dineen

Odds and ends ...

- What is the distribution of the number of heads in n tosses of a biased coin?
- Under what conditions does the Binomial distribution apply to samples taken without replacement from a finite population? When interested in assessing the distribution of a R.V., X , the number of observations in the sample (of n) with one specific characteristic, where $n/N < 0.1$ and a proportion p have the characteristic of interest in the beginning of the experiment.

Slide 21 STAT 13, UCLA, Joe Dineen

Binomial Probabilities – the moment we all have been waiting for!

- Suppose $X \sim \text{Binomial}(n, p)$, then the **probability**

$$P(X = x) = \binom{n}{x} p^x (1-p)^{(n-x)}, \quad 0 \leq x \leq n$$

- Where the **binomial coefficients** are defined by

$$\binom{n}{x} = \frac{n!}{(n-x)! x!}, \quad n! = 1 \times 2 \times 3 \times \dots \times (n-1) \times n$$

Slide 22 STAT 13, UCLA, Joe Dineen

Binomial Formula with examples

- Does the Binomial **probability** satisfy the requirements?

$$\sum_x P(X = x) = \sum_x \binom{n}{x} p^x (1-p)^{(n-x)} = (p + (1-p))^n = 1$$

- Explicit examples for $n=2$, do the case $n=3$ at home!

$$\sum_{x=0}^2 \binom{2}{x} p^x (1-p)^{(2-x)} = \left\{ \begin{array}{l} \text{Three terms in the sum} \\ \binom{2}{0} p^0 (1-p)^2 + \binom{2}{1} p^1 (1-p)^1 + \binom{2}{2} p^2 (1-p)^0 \\ 1 \times 1 \times (1-p)^2 + 2 \times p \times (1-p) + 1 \times p^2 \times 1 = \\ (p + (1-p))^2 = 1 \end{array} \right. \left\{ \begin{array}{l} \text{Usual} \\ \text{quadratic-} \\ \text{expansion} \\ \text{formula} \end{array} \right.$$

Slide 23 STAT 13, UCLA, Joe Dineen

Expected values

- The game of chance: cost to play: \$1.50; Prices {\$1, \$2, \$3}, probabilities of winning each price are {0.6, 0.3, 0.1}, respectively.
- **Should** we play the game? What are our **chances** of winning/loosing?

Prize (\$)	x	1	2	3	
Probability	pr(x)	0.6	0.3	0.1	
What we would "expect" from 100 games					
Number of games won		0.6 × 100	0.3 × 100	0.1 × 100	add across row
\$ won		1 × 0.6 × 100	2 × 0.3 × 100	3 × 0.1 × 100	Sum
Total prize money = Sum:		Average prize money = Sum/100 = 1 × 0.6 + 2 × 0.3 + 3 × 0.1 = 1.5			

Theoretically Fair Game: price to play EQ the expected return!

Slide 24 STAT 13, UCLA, Joe Dineen

TABLE 5.4.1 Average Winnings from a Game conducted N times

Number of games played (N)	Prize won in dollars(x)			Average winnings per game (\bar{x})
	1	2	3	
100	64 (.64)	25 (.25)	11 (.11)	1.7
1,000	573 (.573)	316 (.316)	111 (.111)	1.538
10,000	5995 (.5995)	3015 (.3015)	990 (.099)	1.4995
20,000	11917 (.5959)	6080 (.3040)	2000 (.1001)	1.5042
30,000	17946 (.5982)	9049 (.3016)	3005 (.1002)	1.5020
∞	(.6)	(.3)	(.1)	1.5

So far we looked at the theoretical expectation of the game. Now we simulate the game on a computer to obtain random samples from our distribution, according to the probabilities {0.6, 0.3, 0.1}.

Slide 25 STAT 13, UCLA, Joe Dineen

Definition of the **expected value**, in general.

- The expected value:

$$E(X) = \sum_{\text{all } x} x P(x) \left(= \int x P(x) dx \right)_{\text{all } X}$$

- = Sum of (value times probability of value)

Slide 26 STAT 13, UCLA, Joe Dineen

Example

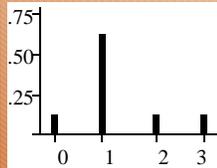
In the at least one of each or at most 3 children example, where $X = \{\text{number of Girls}\}$ we have:

X	0	1	2	3
$\text{pr}(x)$	$\frac{1}{8}$	$\frac{5}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

$$E(X) = \sum_x xP(x)$$

$$= 0 \times \frac{1}{8} + 1 \times \frac{5}{8} + 2 \times \frac{1}{8} + 3 \times \frac{1}{8}$$

$$= 1.25$$



Slide 27 STAT 13, UCLA, Joe Dineen

The expected value and population mean

$\mu_x = E(X)$ is called the *mean* of the distribution of X .

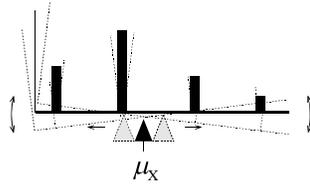
$\mu_x = E(X)$ is usually called the *population mean*.

μ_x is the point where the bar graph of $P(X = x)$ balances.

Slide 28 STAT 13, UCLA, Joe Dineen

The expected value as the point of balance

$\text{pr}(x)$



The mean μ_x is the balance point.

Slide 29 STAT 13, UCLA, Joe Dineen

Population standard deviation

The *population standard deviation* is

$$\text{sd}(X) = \sqrt{E[(X - \mu)^2]}$$

Note that if X is a RV, then $(X - \mu)$ is also a RV, and so is $(X - \mu)^2$. Hence, the *expectation*, $E[(X - \mu)^2]$, makes sense.

Slide 30 STAT 13, UCLA, Joe Dineen

Population mean & standard deviation

Expected value:

$$E(X) = \sum_x xP(X = x)$$

Variance

$$\text{Var}(X) = \sum_x (x - E(x))^2 P(X = x)$$

Standard Deviation

$$\text{SD}(X) = \sqrt{\text{Var}(X)} = \sqrt{\sum_x (x - E(x))^2 P(X = x)}$$

Slide 31 STAT 13, UCLA, Joe Dineen

For the Binomial distribution . . . mean

$$E(X) = np,$$

$$\text{sd}(X) = \sqrt{np(1-p)}$$

$$E(X) = \sum_{x=0}^n x \binom{n}{x} p^x (1-p)^{n-x} =$$

$E(X) = \text{Sum}(\text{Value} \times \text{Probability})$

$$\sum_{x=1}^n x \binom{n}{x} p^x (1-p)^{n-x} =$$

If $x=0$, the entire term is zero

$$\sum_{x=0}^{n-1} (x+1) \binom{n}{x+1} p^{x+1} (1-p)^{n-1-x} =$$

Change variables: $x \rightarrow (x+1)$

Slide 32 STAT 13, UCLA, Joe Dineen

Linear Scaling (affine transformations) $aX + b$

For any constants a and b , the **expectation of the RV $aX + b$** is equal to the sum of the product of a and the expectation of the RV X and the constant b .

$$E(aX + b) = a E(X) + b$$

And similarly for the standard deviation (b , an additive factor, does not affect the SD):

$$SD(aX + b) = |a| SD(X)$$

Slide 38 STAT 13, UCLA, Joe Dineen

Linear Scaling (affine transformations) $aX + b$

Why is that so?

$$E(aX + b) = a E(X) + b \quad SD(aX + b) = |a| SD(X)$$

$$E(aX + b) = \sum_{x=0}^n (a x + b) P(X = x) =$$

$$\sum_{x=0}^n a x P(X = x) + \sum_{x=0}^n b P(X = x) =$$

$$a \sum_{x=0}^n x P(X = x) + b \sum_{x=0}^n P(X = x) =$$

$$aE(X) + b \times 1 = aE(X) + b.$$

Slide 39 STAT 13, UCLA, Joe Dineen

Linear Scaling (affine transformations) $aX + b$

And why do we care?

$$E(aX + b) = a E(X) + b \quad SD(aX + b) = |a| SD(X)$$

-E.g., say the rules for the game of chance we saw before change and the new pay-off is as follows: $\{\$0, \$1.50, \$3\}$, with **probabilities** of $\{0.6, 0.3, 0.1\}$, as before. What is the newly expected return of the game? Remember the old expectation was equal to the entrance fee of \$1.50, and the game was **fair**!

$$Y = 3(X-1)/2$$

$$\{\$1, \$2, \$3\} \rightarrow \{\$0, \$1.50, \$3\},$$

$$E(Y) = 3/2 E(X) - 3/2 = 3 / 4 = \$0.75$$

And the game became clearly biased. Note how easy it is to compute $E(Y)$.

Slide 41 STAT 13, UCLA, Joe Dineen

Review

- What does the expected value of X tell you about? (Expected outcome from an experiment regarding the characteristics measured by the RV X)
- Why is the **expected value** also called the **population mean**? [because for finite population $E(X)$ is the ordinary mean (average)]
- What is the relationship between the population mean and the bar graph of the probability function? (balances the graph)
- What are the mean and standard deviation of the Binomial distribution? (np ; $np(1-p)$)
- Why is $SD(X+10) = SD(X)$?
- Why is $SD(2X) = 2SD(X)$? (Section 5.4.3)

Slide 42 STAT 13, UCLA, Joe Dineen