Ch. 4: Discrete Random Variables

A random variable: is a variable that describes the possible outcomes of a statistical experiment

Example: We toss a coin 3 times and record the number of times we get heads.

Define the Random Variable:

X = the number of times (out of 3) that we get heads

Values of the Variable: 0, 1, 2, 3

* We use capital letters (X) to describe the variable (in words like in Ch. 1) + lower case letters (x=1) to describe values (data, numbers)

A random variable can be either continuous or discrete.

- In Ch. 4, we will talk about discrete random variables
- In Ch. 5 – 11, we will talk about continuous random variables

Probability Distribution Function (PDF)

is a function that assigns a unique probability to every possible value of a random variable

PDF Probabilities have two characteristics:

1. For each value x, $0 \le P(x) \le 1$

2. Sum of the probabilities is 1

Notation for a PDF probability:

$$P(X=x) \quad \text{or} \quad P(x)$$

Example: I toss a coin 3 times and observe the number of times I get heads

1. Write the R.V.

   X = number of heads

2. Write the sample space S:

   {HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}

3. Construct the PDF

x	P(x)
0	1/8
1	3/8
2	3/8
3	1/8

This is the PDF

Expected Value of an Experiment μ:

This is the long term average of the experiment

μ = sum of all $x \cdot P(x)$

x	P(x)
0	$\frac{1}{8}$
1	$\frac{3}{8}$
2	$\frac{3}{8}$
3	$\frac{1}{8}$

So in the long run, on average we expect to get 1.5 heads when we flip 3 coins

μ=expected value=

$$0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{3}{8} + \frac{6}{8} + \frac{3}{8}$$

$$= \frac{12}{8} = 1.5$$

Law of Large Numbers:

As we do an experiment more and more times, the experimental average will get closer to μ

4.10 Practice 1: Discrete Distribution[10]

4.10.1 Student Learning Outcomes

- The student will analyze the properties of a discrete distribution.

4.10.2 Given:

A ballet instructor is interested in knowing what percent of each year's class will continue on to the next, so that she can plan what classes to offer. Over the years, she has established the following probability distribution.

- Let X = the number of years a student will study ballet with the teacher.
- Let $P(x)$ = the probability that a student will study ballet x years.

4.10.3 Organize the Data

Complete the table below using the data provided.

x	P(x)	x*P(x)
1	0.10	0.10
2	0.05	0.10
3	0.10	0.30
4	0.15	0.60
5	0.30	1.50
6	0.20	1.20
7	0.10	0.70

Table 4.9

1   4.5 = μ = expected value

Exercise 4.10.1
In words, define the Random Variable X.

Exercise 4.10.2
$P(x = 4) =$ 0.15

Exercise 4.10.3
$P(x < 4) =$ 0.25

Exercise 4.10.4
On average, how many years would you expect a child to study ballet with this teacher?

4.5 yrs.

4.10.4 Discussion Question

Exercise 4.10.5
What does the column "P(x)" sum to and why?

Exercise 4.10.6
What does the column "$x * $P(x)" sum to and why?

[10]This content is available online at <http://cnx.org/content/m16830/1.14/>.

Binomial Distribution or Binomial PDF

In order to get a binomial distribution, we must start out with a binomial experiment

*Binomial Experiment:

1. The same activity is repeated a fixed number of times

   Examples:
   - Toss a coin 10 times (=n)
   - Roll 3 dice 7 times (=n)
   - Ask 8 (=n) people out of 50 the same question

   Each repetition is called a trial

   n = number of trials

2. **Each trial has only 2 possible outcomes:**

   success **or** failure

   p = probability of success

   q = probability of failure

   p + q = 1

3. **The trials must be independent**

In each trial, p and q remain the same

2

Example: Suppose I roll a single die 5 times. I am interested in how many times I roll a 6.

Trial: One roll of the die n = 5

success = 6 p = $\frac{1}{6}$

failure = 1, 2, 3, 4, 5 q = $\frac{5}{6}$

Are the trials independent? Why?

Yes. . .

The outcomes of a binomial experiment fit a Binomial Distribution.

We write: **X~B(n,p)** : We say "X has a binomial distribution with parameters n and p"

n and p are called parameters

μ = mean = expected value = $n \cdot p$

σ = standard deviation = $\sqrt{npq}$

Chapter 4: Binomial Experiments Practice

1. We roll 6 different colored dice. Of interest is the number of dice that show a "1"

Trial: roll of 1 die. n = 6

Success = 1 p = $\frac{1}{6}$

Failure = 2, 3, 4, 5, 6 q = $\frac{5}{6}$

Are trials independent? Yes

2. About 85% of graduating De Anza students attend their graduation. We randomly survey 22 graduating De Anza students this quarter and ask them if they are going to attend the graduation ceremony.

Trial: Asking a DA if they're going to graduation n = 22

Success = Yes p = $0.85 = \frac{85}{100} = 85\%$

Failure = No q = $0.15 = \frac{15}{100} = 15\%$

Are trials independent? Yes

3. About 3% of all fortune cookies contain an extra fortune. We buy a bag of 144 fortune cookies. We are interested in how many of them contain an extra fortune inside.

Trial: opening 1 cookie n = 144

Success = more than 1 fortune p = $3\% = 0.03$

Failure = 1 fortune (or 0) q = $97\% = 0.97$

Are trials independent? Yes

4

p.85

Example: According to a recent article, 60% of U.S. adults say they experienced a sleep problem in the last year. 8 U.S. adults are randomly surveyed. We are interested in the number who say that they experienced a sleep problem in the last year.

a. Define the random variable in words.

X = number (out of 8) who experienced a sleep problem

b. $X \sim B(8, 0.6)$ — 8 ↑ n, 0.6 ↑ p

① 8 ← n trials
② success = yes, p = 0.6
③ indep ✓

c. How may people out of 8 do we expect → μ to have experienced a sleep problem? What is the standard deviation? ← σ

$\mu = np = 8(0.6) = 4.8$ people

$\sigma = \sqrt{8(0.6)(0.4)} \approx 1.3656$

d. Calculate the PDF.

P(X=0) → DISTR → binompdf(n, p, x)
" (8, 0.6, 0)

x	P(x) P.F	P(X ≤ x) C.P.F
0	0.0007	0.0007
1	0.0079	0.0085
2		
3		
4		
5		
6		
7		
8		

Ck total = 1

DISTR → binomcdf(8, 0.6, 0)

d. Calculate the PDF.

x	P(X = x)	P(X ≤ x)
0	0.0007	0.0007
1	0.0079	0.0085
2	0.0413	0.0498
3	0.1239	0.1737
4	0.2322	0.4059
5	0.2787	0.6846
6	0.2090	0.8936
7	0.0896	0.9832
8	0.0168	1.0000

e. What is the probability that exactly 4 will have experienced a sleep problem?

$P(X=4) = 0.2322$

f. What is the probability that at most 6 will have experienced a sleep problem?

$P(X \leq 6) = 0.8936$

g. What is the probability that more than 6 will have experienced a sleep problem?

$P(X>6) = 1 - P(X \leq 6) = 0.1064$

$= P(X=7 \text{ or } X=8) = 0.0896 + 0.0168$

$= 0.1064$

h. What is the probability that at least 6 will have experienced a sleep problem?

$P(X \geq 6) = P(X=6 \text{ or } X=7 \text{ or } X=8)$

$= P(X>5) = 1 - P(X \leq 5)$

$= 0.3154$

Translating Expressions

X is at most 4	$X \leq 4$	$P(X \leq x) = \text{binomcdf}(n,p,x)$
X is less than 4	$X < 4$	$P(X < x) = P(X \leq x-1)$ $P(X < 4) = P(X \leq 3)$
X is at least 4	$X \geq 4$	$P(X \geq x) = 1 - P(X \leq x-1)$ $P(X \geq 4) = 1 - P(X \leq 3)$
X is more than 4	$X > 4$	$P(X > x) = 1 - P(X \leq x)$ $P(X > 4) = 1 - P(X \leq 4)$