Lecture 6: Discrete Random Variables

Probability Theory and Applications

Fall 2005

September 20

Random Variables (RVs)

X is a discrete a random variable (RV) if its set of possible values (value set) is finite or countably infinite.

A discrete probability density function (pdf) of X is a list of formula that show the possible values of X and the corresponding probabilities.

$$f(x) = P[X = x] \qquad x = x_1, x_2, \ldots$$

Example

Toss 3 fair coins

Let random variable

Y= # of heads

Value Set ={ 0,1,2,3}

y	0	1	2	3
$f(y)$	$1/8$	$3/8$	$3/8$	$1/8$

Example

Toss 3 fair coins

Let random variable

X= # of heads - # of tails

Value Set ={ -3,-1,1,3 }

x	-3	-1	1	3
$f(x)$	$1/8$	$3/8$	$3/8$	$1/8$

Tosses	X
TTT	-3
HTT	-1
HHT	1
HHH	3
THT	-1
TTH	-1
HTH	1
THH	1

Theorem

A function f(x) is a discrete pdf iff it satisfies the following properties for at most a countably infinite set of reals $x_1, x_2, \ldots.$

$$f(x) \geq 0 \text{ and } \sum_{all\ x_i} f(x_i) = 1.$$

Note $1 \geq f(x) \geq 0$ $$f(x) := P(X = x)$$

and *f(x)=0* if *x* is not in the value set

A. Given pdf, calculate probabilities

Using toss 3 fair coin experiments and

X=# heads -# of tails

What is probability X<0,

$P(X<0)=P(X=-3)+P(X=-1)$

$=f(-3)+f(-1)=1/3+3/8=1/2$

B. Given problem, find pdf

Experiment: Toss a fair coin until head shows up. Let Y=# of tosses

S	Y	f(y)

Given problem, find pdf

Experiment: Toss a fair coin until head shows up. Let Y=# of tosses

S	Y	p(y)
H	1	½
TH	2	$(½)^2$
TTH	3	$(½)^3$
TTTH	4	$(½)^4$
…..		
	y	$(½)^y$

PDF

$$recall$$
$$\sum_{i=0}^{\infty}(1-a)^{-1}=\frac{1}{1-a} \quad 0<a<1$$

- The pdf is

$$f(x)=\left(\frac{1}{2}\right)^{x} \qquad x=1,2,3,4,\ldots$$

- Check

$$\sum_{x=1}^{\infty}f(x)=\sum_{x=1}^{\infty}\left(\frac{1}{2}\right)^{x}=\sum_{x=0}^{\infty}\left(\frac{1}{2}\right)^{x}-1=2-1=1$$

- P(head occurs < 3 tosses) = P(X<3)

  = P(X=1)+P(X=2)=1/2+1/4=3/4

C. Find constant in pdf

Given X has pdf

$$f(x) = cx^2 \quad x = 1, 2, \ldots, 5$$

Find c

$$1 = \sum_{x=1}^{5} f(x) = c\left[1 + 4 + ... + 25\right]$$

$$1 = 55c$$

$$c = \frac{1}{55}$$

So real pdf is

$$f(x) = \frac{x^2}{55} \quad x = 1, 2, \ldots, 5$$

$$P(X = 4) = \frac{16}{55}$$

P(X<=3)=P(X=1)+P(X=2)+P(X=3)

Cummulative Distribution Function (cdf)

The distribution of a RV X can be defined using the **cdf**

$$F(x) = P(X \le x) \qquad cdf$$

$$f(x) = P(X = x) \qquad pdf$$

CDF of last example

F(x)	x
0	x<1
1/55	1≤x<2
5/55	2 ≤ x<3
14/55	3 ≤ x<4
30/55	4 ≤ x<5
55/55	5 ≤ x

- Sketch it:

Calculate from cdf

- $P(X \leq 3.5) = F(3.5) = 14/55$
- $P(X=3) = P(X<=3) - P(X<3)$
  $= 14/55 - 4/55 = 9/55$
- $P(X \leq 4 | x > 1) = 29/54$

Standard Discrete RV

- Now
  - Definition and uses in problems
- Later
  - Theoretical Properties

Bernoulli Trial

- Experiment with only two possible outcomes

  s= success f=failure
- For each trial, the probability of success is fixed to p: $P(s)=p$ $P(f)=1-p=q$
- Each trial is independent

Examples:

Flip coin.

Choose a red ball from an urn with replacement.

Roll a die and observe if value is 6 or not.

Binomial Distribution

The RV *X* is the number of success in n Bernoulli trials.

Parameters:

p=prob. of success   n =# of trials

p.d.f.

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, \ldots, n$$

Binomial Example: HW 1

X=# of heads in 3 tosses

Success is head

n=3 p=1/2

$$f(0)=\binom{3}{0}\left(\frac{1}{2}\right)^{0}\left(\frac{1}{2}\right)^{3-0}=\frac{1}{8} \qquad f(1)=\binom{3}{1}\left(\frac{1}{2}\right)^{1}\left(\frac{1}{2}\right)^{2}=\frac{3}{8}$$

$$f(2)=\binom{3}{2}\left(\frac{1}{2}\right)^{2}\left(\frac{1}{2}\right)^{1}=\frac{3}{8} \qquad f(0)=\binom{3}{3}\left(\frac{1}{2}\right)^{3}\left(\frac{1}{2}\right)^{0}=\frac{1}{8}$$

Binomial Example 4

What is prob that at most 3 of 10 computer chips in a lot will be bad, if the prob that any one chip is bad is 0.05?

$$P(Y \le 3) = P(Y = 0) + P(Y = 1) + P(Y = 2) + P(Y = 3)$$
$$= f(0) + f(1) + f(2) + f(3) = .9945$$

This is just c.d.f F(3). Use table for cdf of binomial in Appendix c in book.
Say we pick 20. Y~binomial(n=20 p= .05)
P(Y<=5)=.9997 P(2<Y<=5)=F(5)-F(2)=.9997-.9245)

Quick Experiment

Goto

http://www.math.uah.edu/stat/applets/BinomialCoinExperiment.xhtml

Do experiment for n=50

What happens as p changes?

Binomial Example 4

According to *Chemical Engineering Progress* Nov 1990, approx 30% of all paperwork failures are caused by operator error (OE).

1. What is prob that in the next 20 paperwork failures, at least 10 are caused by OE?
2. What is probability that no more than 4 of 20 are due to operator error?
3. Suppose at a particular plant, exactly 8 of the last 20 are operational errors. Does it seem that the 30% figure applies to this plant?

Sampling with replacement = Binomial

Experiment

M defectives

N-M goods

N total

→ Draw n items with replacement

Let X=number of defectives
n=number of trials
p = M/N

X~binomial(n,p)

Sampling without replacement = Hypergeometric

Experiment

Draw n items without replacement

M defectives

N-M goods

N total

Let X=number of defectives
n=number of trials
M=number of items of interest
N=total

X~hypergeometric (n,M,N)

$$f(x) = P(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

Hypergeometric

Of 120 people applying for a job, only 80 are qualified. If 5 people are picked at random, what is prob that exactly two will be qualified to get a job?

$X = \#$ of qualified applicants

X is hypergeometric with $m = 80, N = 120, n = 5$

$$f(2) = P(X = 2) = \frac{\binom{80}{2}\binom{40}{3}}{\binom{120}{5}} = 0.164$$

Binomial Approximation of Hypergeometric

Assume people are picked with replacement

$X = \#$ of qualified applicants

X is approx. binomial $p = 80/120 \quad n = 5$

$$f(2) = P(X = 2) = \binom{5}{2}(80/120)^2(40/120)^3 = 0.165$$

Problem

300 people are selected for jury duty. 30 of 300 are under 25. 12 jurors are selected are all over 25. Defendant is convicted.

In appeal, defense attorney argues that jury selection was unfair because all the jurors are over 25. If jurors are selected at random, what is the probability that at least one will be under 25?

Problem

Quality control engineer wants to check if 95% of components shipped by company are flawless. She selects 20 from very large lot. Lot passes if all are flawless. Checked otherwise. Find prob.

1. Hold a lot for complete check even though 95% are flawless.
2. Passing a lot when only 90% (80%, 70%) are flawless.