11 BINOMIAL DISTRIBUTION

Examples:

1. Five terminals on an on-line computer system are attached to a communication line to the central computer system. The probability that any terminal is ready to transmit is 0.95.
   Let X denote the number of ready terminals.

2. A fair coin is tossed 10 times ; success and failure are "heads" and "tails" respectively, each with probability, .5.
   Let X be the number of heads (successes) obtained.

3. It is known that 20% of integrated circuit chips on a production line are defective. To maintain and monitor the quality of the chips, a sample of twenty chips is selected at regular intervals for inspection.
   Let X denote the number of defectives found in the sample.

4. It is known that 1% of bits transmitted through a digital transmission are received in error. One hundred bits are transmitted each day.
   Let X denote the number of bits found in error each day.

BINOMIAL CONDITIONS

1. An experiment consists of n repeated trials.
2. Each trial has two possible outcomes: success or failure.
3. The probability of a success p is constant from trial to trial.
4. Repeated trials are independent.

Let $X =$ number of successes in n trials

X is a BINOMIAL random variable.

General Binomial Distribution

$n =$ no of trials

$p =$ probability of success

$q = 1 - p =$ probability of failure

X $=$ no of successes in n trials

$$P(X = x) = \binom{n}{x} p^x q^{n-x}$$

In Example 1, we might ask the probability that 3 ready terminals in 5 terminals:

$$P(X = 3) = \binom{5}{3} (.95)^3 (.05)^2$$

Calculating binomial pdfs with *R*

Example 1, the probability of getting exactly 3 ready terminals in 5:

```
dbinom(x = 3,  size =5, prob = .95)
[1] 0.02143438
```

For all the probabilities

```
x<- 0:5
```

and

```
dbinom(x, size= 5, prob =.95)
[1] 0.0000003125 0.0000296875 0.0011281250 0.0214343750
[5] 0.2036265625, 0.773780937
```

Example 2, tossing a coin 10 times

```
round(dbinom(x, 10, .5), 4)
 [1] 0.0010 0.0098 0.0439 0.1172 0.2051 0.2461 0.2051 0.1172 0
[11] 0.0010
```

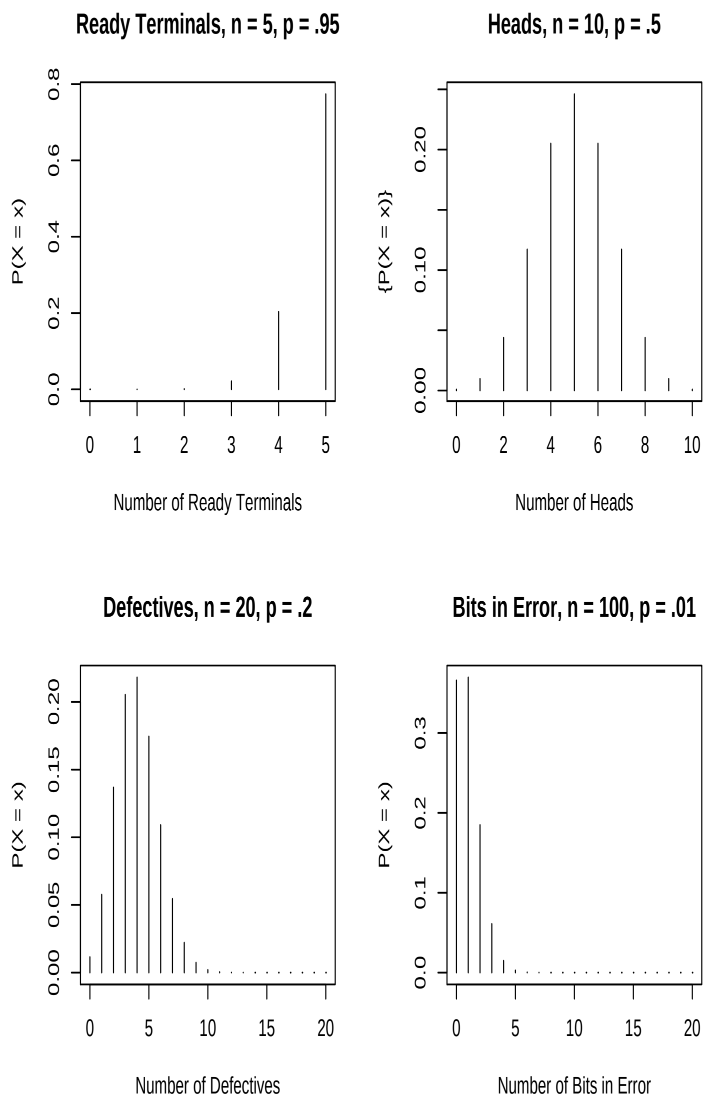
Examples 3 and 4.

```
dbinom(x, 20,  .2)
```

and

```
dbinom(x,100, .01)
```

Binomial Probability Density Functions



Plotting binomial pdfs in *R*

```
par(mfrow = c(2,2)) # multiframe

x<-0:5 #Example 11.1
plot(x, dbinom(x, size = 5, prob = .95),
 xlab = "Number of Ready Terminals",
 ylab = "P(X = x)",  type = "h",
 main = "Ready Terminals, n = 5, p = .95")

x<-0:10 #Example 11.2
plot(x, dbinom(x, size = 10, prob = .5),
 xlab = "Number of Heads",
 ylab = "P(X = x)",  type = "h",
 main = "Heads, n = 10, p = .5")

x<-0:20 #Example 11.3
plot(x, dbinom(x, size = 20, prob = .2),
 xlab = "Number of Defectives",
 ylab = "P(X = x)", type = "h",
 main = "Defectives, n = 20, p = .2")

x<-0:20 #Example 11.4. No need for full range here
plot(x, dbinom(x, size = 100, prob = .01),
 xlab = "Number of Bits in Error",
 ylab = "P(X = x)",  type = "h",
 main = "Bits in Error, n = 100, p = .01")
```

Cumulative Distribution Function

$$\begin{aligned} P(X \leq x) &= P(X=0) + P(X=1) + \cdots + P(X=x) \\ &= q^n + \binom{n}{1} p^1 \cdot q^{n-1} + \cdots + \binom{n}{x} p^x \cdot q^{n-x} \end{aligned}$$

Example 1:

P(less than or equal to 3 terminals will be ready)

$$\begin{aligned} &= P(X=0) + P(X=1) + P(X=2) + P(X=3) \\ &= (.05)^5 + \binom{5}{1}(.95)^1(.05)^4 + \binom{5}{2}(.95)^2(.05)^3 + \binom{5}{3}(.95)^3(.05)^1 \\ &= .0000003 + .00003 + .00113 + .02143 \\ &\approx .0226 \end{aligned}$$

Also

$$\begin{aligned} &= P(\text{4 or more terminals will be ready}) \\ &= P(X \geq 4) \\ &= 1 - P(X \leq 3) \\ &\approx 1 - .0226 \\ &= .9774 \end{aligned}$$

Calculating binomial cdfs in *R*

Example 1: n = 5, p = .95

P(X $\leq$ 3):

```
pbinom(3, 5, .95)
[1] 0.0225925
```

P(X $>$ 3):

```
1-pbinom(3, 5, .95)
[1] 0.9774075
```

Example 3: n = 20, p = .2

P(X $\leq$ 4):

```
pbinom (4, size = 20, prob =.2)
0.6296483
```

P(X $>$ 4) = 1- P(X $\leq$ 4)

```
1- pbinom(4, size = 20, prob = .2)
[1] 0.3703517
```

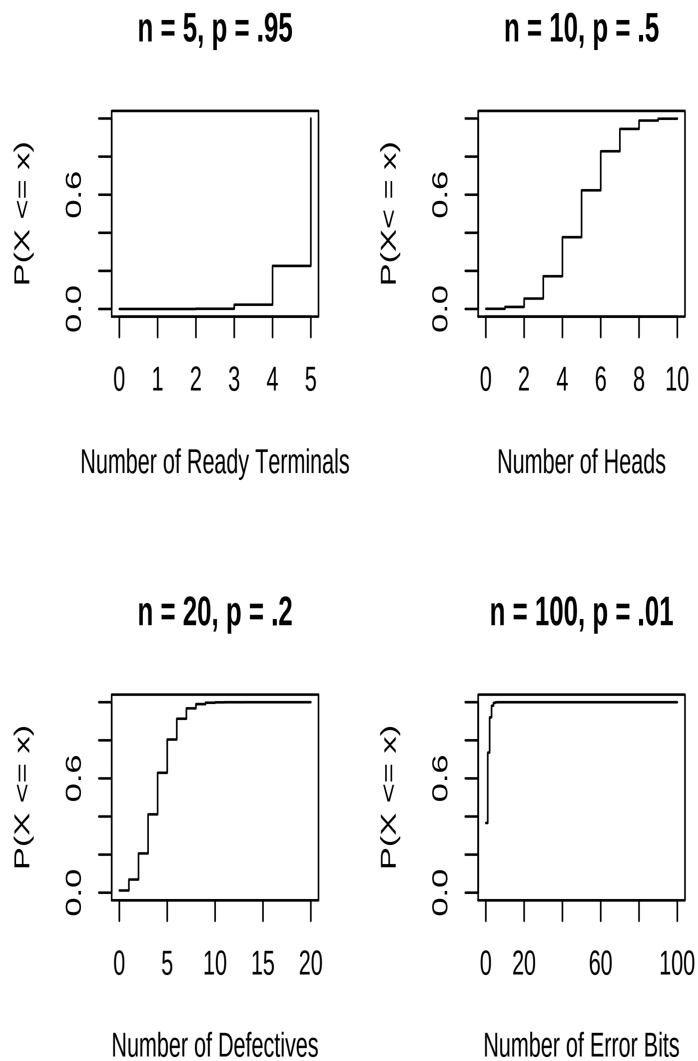
For all cumulative probabilities

```
x<-0:20
prob<- pbinom(x, size = 20, prob= .2)
```

round to 4 decimal places:

```
round(prob, 4)
 [1] 0.0115 0.0692 0.2061 0.4114 0.6296
```

Binomial Cumulative Distribution Functions



11.1 Plotting Binomial cdfs

```
par(mfrow = c(2,2)) #multiframe
x<-0:5
plot(x, pbinom(x, size = 5, prob = .95),
 xlab = "Number of Ready Terminals",
 ylab = "P(X <= x)",  ylim = c(0, 1),
 type = "s", main = "n = 5, p = .95")

x<-0:10
plot(x, pbinom(x, size = 10, prob = .5),
 xlab = "Number of Heads",
 ylab = "P(X< = x)",  ylim = c(0, 1),
 type = "s", main=  "n = 10, p = .5")

x<-0:20
plot(x, pbinom(x, size = 20, prob = .2),
 xlab = "Number of Defectives",
 ylab = "P(X <= x)", ylim = c(0, 1),
 type = "s",  main = "n = 20, p = .2")

x<-0:100
plot(x, pbinom(x, size = 100, prob = .01),
 xlab = "Number of Error Bits",
 ylab = "P(X <= x)",  ylim = c(0, 1),
 type = "s", main = "n = 100, p = .01")
```

The Quantile Function

Example 3, the batches of integrated chips, we might ask:

Up to how many defectives will the batches contain with at least 95% certainty?

Choose k

$$P(X \leq k) \geq .95$$

```
qbinom(.95, 20, .2)
[1] 7
```

95% of these batches would contain less than or equal to 7 defectives.

Example

Suppose there are n frames per packet and the probability that a frame gets through without an error is $.999$.

What is the maximum size that a packet can be so that the probability that it contains no frame in error is at least .9 ?

Solution

X =the number of frames in error in a packet of size n.

$$P(X = 0) = .999^n$$

Choose n so that

$$.999^n > 0.9$$

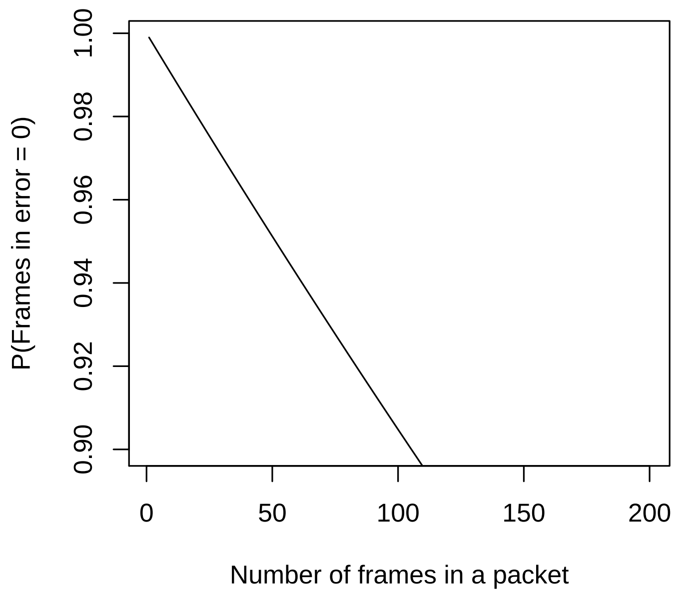
n = 106

When the packet size reaches 106, the probability of it being error free is below 90%

Check in *R*

```
> qbinom(.9, 105, .001)
[1] 0
> qbinom(.9, 106, .001)
[1] 1
```

Error Free Packets



R code:

```
n <- 1: 200
plot(n, .999^n,  type = "l",
   ylab = "P(Frames in error = 0)",
    xlab="Number of frames in a packet",
    ylim= c(.9, .999))
```

Machine Learning

Individual decisions of an ensemble of classifiers are combined to classify new examples in order to improve the classification accuracy.

Example:

3 classifiers used to classify a new example, each having a probability p = .7 of correctly classifying a new case.

Calculate the probability that the new case will be correctly classified if a majority decision is made.

Solution:
X = number of correct classifications with 3 classifiers.
X is binomial with $n = 3$ and $p = .7$.

A majority vote means $X \geq 2$.

The probability of correctly classifying a new example based on a majority decision:

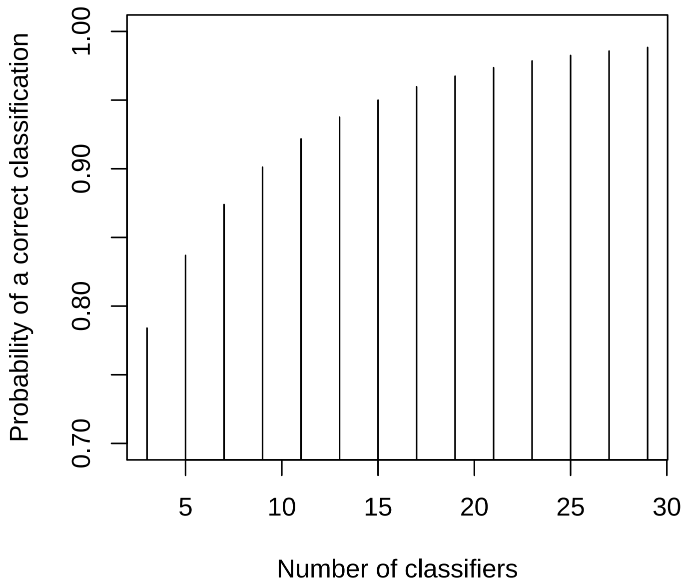
$$\begin{aligned} P(X \geq 2) &= P(X = 2) + P(X = 3) \\ &= \binom{3}{2} .7^2 .3 + .7^3 \end{aligned}$$

```
1-pbinom(1, 3, .7)
0.784
```

With 5 classifiers and a majority decision of 3:

```
1-pbinom(2, 5, .7)
[1] 0.83692
```

Increasing the number of classifiers when p = .7



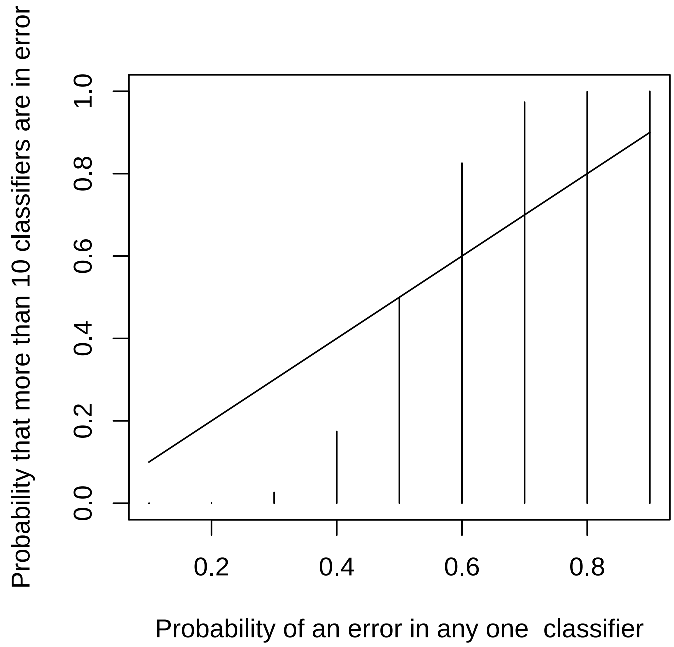
R code:

```
n <- seq(3, 29, 2)
majority <-(n+1)/2

probmajority <- 1-pbinom(majority-1, n, .7)

plot(n, probmajority,
 xlab = "Number of classifiers",
 ylab = "Probability of a correct classification",
 ylim = c(.7, 1), type = "h")
```

Classification accuracy with 21 classifiers for varying p



When error probability $p < .5$, classification accuracy improves

When error probability $p > .5$, classification accuracy worstens.

Key to successful ensemble methods is to construct individual classifiers with error rates below .5.

R code

```
p <- c(.1, .2, .3, .4, .5, .6, .7, .8, .9)

plot(p, 1-pbinom(10, 21, p),
xlab = "Probability of an error in any one  classifier",
ylab = "Probability that more than 10 classifiers are in error",
type =  "h")

lines(p,  p)
```

Binomial Expectations

1. In a laboratory with 5 terminals, how many terminals are ready to transmit on average, and how will this vary from day to day?

2. How many heads on average do we get when we toss a coin 10 times, and how will this vary from set to set?

3. How many defectives on average do we get in samples of size 20, and how will this vary from sample to sample?

4. How many bits in error would we expect in the 100 daily transmissions, and how will the number of errors vary from day to day?

Mean of the Binomial Distribution

$$\begin{aligned}\mu = E(X) &= \sum_{x=0}^{n} x p(x) \\ &= \sum_{x=0}^{n} x \binom{n}{x} p^x \cdot q^{n-x}.\end{aligned}$$

Example

From past experience it is known that there is a 25% chance that a source program written by a certain programmer compiles successfully. Each day the programmer writes five programs.

The probabilities of x programs compiling each day

$$P(X = x) = \binom{5}{x} .25^x .75^{5-x} \text{for} \quad x = 0, 1, 2, \cdots, 5$$

Binomial Probabilities with $n = 5$ and $p = .25$

Number that compiles x	0	1	2	3	4	5
Probability $p(x)$	.237	.396	.264	.088	.015	.001

$$E(X) = 0\times.237+1\times.396+2\times.264+3\times.088+4\times.014+5\times.001 = 1.25$$

On average, 1.25 programs compile each day.

Mean of the Binomial Distribution

$$
\begin{aligned}
E(X) &= = \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x} \\
&= \sum_{x=1}^{n} x \binom{n}{x} p^x q^{n-x} \quad \text{since the first term is 0} \\
&= \sum_{x=1}^{n} n \binom{n-1}{x-1} p^x q^{n-x} \\
&= np \sum_{x-1=0}^{n-1} \binom{n-1}{x-1} p^{x-1} q^{(n-1)-(x-1)} \\
&= np\,(p+q)^{n-1} \\
&= np
\end{aligned}
$$

Variance of the Binomial Distribution

$$\begin{aligned}\sigma^2 = V(X) &= \sum_x (x-\mu)^2 p(x) \\ &= \sum_x (x-np)^2 \binom{n}{x} p^x \cdot q^{n-x}\end{aligned}$$

.

$$\begin{aligned}V(X) &= (0-1.25)^2 \times .237 \\ &\quad +(1-1.25)^2 \times .396 \\ &\quad +(2-1.25)^2 \times .264 \\ &\quad +(3-1.25)^3 \times .088 \\ &\quad +(4-1.25)^2 \times .015 \\ &\quad +(5-1.25)^2 \times .001 \\ &= 0.9375\end{aligned}$$

The standard deviation

$$\sigma = \sqrt{0.9375} = \ 0.9683$$

It can be shown that

$$\sigma^2 = \sum_x (x-np)^2 \binom{n}{x} p^x \cdot q^{n-x} = npq$$

Example 1: In the on-line computer system with 5 terminals each with a probability of .95 of being ready, mean and variance of the number of ready terminals

$$\mu = 5 \times .95 = 4.75.$$

$$\sigma^2 = 5 \times .95 \times .05 = 0.2375$$

Example 2: Tossing a coin 10 times, mean and variance of number of heads

$$\mu = 10 \times .5 = 5$$

$$\sigma^2 = 10 \times .5 \times .5 = 2.5$$

Example 3: In the batches of 20 products from a production line with 20% defective, mean and variance of number of defectives per batch:

$$\mu = 20 \times .2 = 4$$

$$\sigma^2 = 20 \times .2 \times .8 = 3.2$$

Example 4: In the digital transmission system , with 1% transmission error, mean and variance of the number of bits in error per day

$$\mu = 100 \times .01 = 1$$

$$\sigma^2 = 100 \times .01 \times .99 = 0.99$$

R Functions for the Binomial Distribution

- dbinom

  ```
  dbinom(x = 4,  size = 20, prob = .2)
  ```

  or

  ```
  dbinom(4,  20, .2)
  ```

  $P(X = 4)$ with $n = 20$ and $p = .2$

- pbinom

  ```
  pbinom(x = 4,  size = 20, prob = .2)
  ```

  or

  ```
  pbinom(4,  20, .2)
  ```

  $P(X \leq 4)$ with $n = 20$ and $p = .2$

- qbinom

  ```
  qbinom(.95, size = 20, prob = .2)
  ```

  or

  ```
  qbinom(.95,  20, .2)
  [1] 7
  ```

  Choose k so that

  $$P(X \leq k) \geq .95.$$