THE BINOMIAL DISTRIBUTION

Consider an experiment with:

Bernoulli trials:

① n repeated INDEPENDENT TRIALS with two possible outcomes (let's tentatively label them "success" and "failure")

② P(success) is CONSTANT at each trial $= p$

Let the random variable of interest be:

r.v. : the number of successes in the n trials

or:

let X = the number of successes in n independent trials with constant probability p of success.

ex/ toss a ~~fair~~ coin 100 times

X = # H's in 100 tosses

ex/ toss a die 6 times

X = # 4's that come up

ex/ Consider 200 independent parts in a circuit with constant probability of failure p in < 20 hours

X = # failures in ~~less than~~ 20 hours

ex/ P(# failures in 20 hours $\leq$ 40) = ?

We'll derive a general expression for the probability of x successes in n independent trials with constant probability p of success at each trial, but first we'll consider a few small values of n and then generalize:

ex/ Consider tossing a die 3 times

define: success = a "4"
failure = not a "4"

So: $n = 3$ $\quad p = \frac{1}{6}$ $\quad (1-p)(=q) = \frac{5}{6}$

Let X = # of "4"s in 3 tosses

Then the probability distribution of X is:

$$P(X=0) = P(FFF) = \frac{5}{6}\cdot\frac{5}{6}\cdot\frac{5}{6} = \left(\frac{5}{6}\right)^3 = \frac{125}{216}$$

$$P(X=1) = P(SFF \text{ or } FSF \text{ or } FFS) = \frac{1}{6}\cdot\frac{5}{6}\cdot\frac{5}{6} + \frac{5}{6}\frac{1}{6}\frac{5}{6} + \frac{5}{6}\frac{5}{6}\frac{1}{6} = \frac{75}{216}$$

$$P(X=2) = P(SSF \text{ or } SFS \text{ or } FSS) = \frac{1}{6}\frac{1}{6}\frac{5}{6} + \frac{1}{6}\frac{5}{6}\frac{1}{6} + \frac{5}{6}\frac{1}{6}\frac{1}{6} = \frac{15}{216}$$

$$P(X=3) = P(SSS) = \frac{1}{6}\cdot\frac{1}{6}\cdot\frac{1}{6} = \left(\frac{1}{6}\right)^3 = \frac{1}{216}$$

$\Sigma = 1$ ✓

ex/ Now consider tossing a die 5 times.
Let X = # 4's in 5 tosses
where still: success = a "4".
failure = not

$$P(X=2) = P(SSFFF \text{ or } SFSFF \text{ or } \ldots)$$
$$= \frac{1}{6}\cdot\frac{1}{6}\cdot\frac{5}{6}\cdot\frac{5}{6}\cdot\frac{5}{6} + \frac{1}{6}\cdot\frac{5}{6}\cdot\frac{1}{6}\cdot\frac{5}{6}\cdot\frac{5}{6} + \ldots$$

and the number of terms in this sum will be the number of different ways of selecting the two tosses to be successes (out of the total of five tosses). This number of terms is $\binom{5}{2}$, and each term is $\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^3$ so:

$$= \binom{5}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^3$$

Now to generalize to n trials, with
P(success on a trial) = p
and X = # successes in n trials

To find $P(X=j)$, note that there are $\binom{n}{j}$ different outcome sequences with j successes and $(n-j)$ failures and each of these sequences has probability of occurrence of $(p)^j(q)^{n-j}$

so:

If X = # successes in n Bernoulli trials with $p = P(\text{success})$ at each trial

$$P(X=j) = \binom{n}{j} p^j q^{n-j} \quad ; \; j = 0, 1, 2, \ldots, n$$

ex/ A marksman takes 10 shots at a target, with $P(\text{hit}) = .1$ at each shot. $n=10$ $p=.1$ $q=.9$

a) $P(\text{exactly 4 hits}) = \binom{10}{4}(.1)^4(.9)^6 = .0116$

b) $P(\text{2 or more hits}) = P(\text{2 hits}) + P(\text{3 hits}) + \ldots + P(\text{10 hits})$

OR, EASIER:

$= 1 - P(\text{0 or 1 hit})$

$= 1 - \left[\binom{10}{0}(.1)^0(.9)^{10} + \binom{10}{1}(.1)^1(.9)^9\right]$

$= .26$

ex/ If 12% of the population is left-handed, what is the probability that a sample of 20 people

$n=20$ $p=.12$ $q=.88$

a) has exactly 2 left-handers?

$= \binom{20}{2}(.12)^2(.88)^{18} = .274$

b) has at most 2 left-handers

$= P(\text{0 or 1 or 2 left-handers})$

$= \binom{20}{0}(.12)^0(.88)^{20} + \binom{20}{1}(.12)^1(.88)^{19} + \binom{20}{2}(.12)^2(.88)^{18}$

$= .563$

MEAN AND STANDARD DEVIATION OF THE BINOMIAL DISTRIBUTION

Let X be a binomially distributed r.v. with $n = \#$ of independent trials and $p = P(\text{success on single trial})$

Then the average (mean) number of successes in n trials is

$$\boxed{\mu = E(X) = np}$$

exA/ In 100 coin tosses with $P(\text{H on a trial}) = .5$,
n p

if $X = \#$ of Heads in 100 tosses, then

average # of Heads in 100 tosses $= E(X) = np = 100(.5) = 50$ Heads

exB/ $X = \#$ of 6's in 10 tosses of a die
$n = 10$ $p = \frac{1}{6}$

$E(X) = \text{av. \# of 6's} = np = 10\left(\frac{1}{6}\right) = \underline{1.67}$ "6"s

Theorem

If $X \sim bin(n, p)$, then $E(X) = np$

Proof:

$$E(X) = \sum_k k P(X=k)$$

$$= \sum_i x_i p_X(x_i)$$

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

$$= \cancel{np \sum_{k=0}^{n}}$$

$$= \sum_{k=1}^{n} \frac{(k)\,n!}{k!\,(n-k)!} p^k (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} \frac{n!}{(k-1)!\,(n-k)!} p^k (1-p)^{n-k}$$

$$= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!\,(n-k)!} p^{k-1} (1-p)^{n-k}$$

Let $k = i+1$ $\quad (i = k-1)$

$$= np \underbrace{\sum_{i=0}^{n-1} \binom{n-1}{i} p^i (1-p)^{(n-1)-i}}_{[p+(1-p)]^{n-1}}$$

$$= np$$

We will prove the following later:

The <u>standard deviation</u> of the number of successes in n trials is

$$\sigma = \sqrt{np(1-p)}$$

$(\sigma^2 = np(1-p) = npq)$

ex A) from previous page

$n = 100 \quad p = .5$

$$\sigma = \sqrt{n(p)(1-p)} = \sqrt{100(.5)(1-.5)} = \underline{\underline{5}}$$

ex B) from previous page

$n = 10 \quad p = \frac{1}{6}$

$$\sigma = \sqrt{np(1-p)} = \sqrt{10\left(\frac{1}{6}\right)\left(1-\frac{1}{6}\right)} = \sqrt{10\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)} = \sqrt{1.3889}$$

$$\doteq \underline{\underline{1.18}}$$