Negative Binomial Distribution.

Consider a sequence of Bernoulli trials in which we wait for k successes, and define a rv:

$$\underline{X} = \text{trial number of the } k^{th} \text{ success.}$$

The value set of $\underline{X}$ is $k, k+1, k+2, \ldots$

(In order to have k successes, there must have been at least k trials).

$\Rightarrow$ The number of trials is not longer fixed in advance; in fact, the number of trials is a rv.

The rv $\underline{X}$ is called a negative binomial rv.

Proposition. The pmf of the negative binomial rv $\underline{X}$ with parameter k = number of successes (S's) and

$$p = P(S)$$

is

$$\left(p(x) = P(\underline{X}=x) =\right) nb(x;\ k, p) = \binom{x-1}{k-1} p^k (1-p)^{x-k}, \qquad x = k, k+1, \ldots$$

Proof:

$$P(\underline{X} = x) = P(x \text{ trials with last trial being a success and } (k-1) \text{ successes in the first } (x-1) \text{ trials})$$

$$= P((k-1) \text{ successes in the first } (x-1) \text{ trials}) \cdot P(S)$$

$$= \binom{x-1}{k-1} p^{k-1} (1-p)^{(x-1)-(k-1)} \cdot p$$

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

* Special case: waiting for the first success.

$(k=1) \Rightarrow$ $\underline{X}$ = trial number of the first success

= geometric r.v.

$$nb(x; 1, p) = p\,(1-p)^{x-1}, \qquad x = 1, 2, 3 \ldots$$

= geometric distribution.

* To verify that $nb(x; k, p)$ is a proper p.f, we must show that

$$\sum_{x=k}^{\infty} nb(x; k, p) = \underline{1}.$$

$\Rightarrow$ That is; to show:

$$\sum_{x=k}^{\infty} \binom{x-1}{k-1} p^{k} (1-p)^{x-k} = \underline{1}$$

* <u>Negative binomial series:</u>

for $|u| < 1$,
$$(1-u)^{-n} = \sum_{j=0}^{\infty} \binom{n+j-1}{j} u^{j}$$

$$= 1 + nu + \frac{n(n+1)}{2!} u^2 + \cdots$$

- Taking $j = x - k$ = the number of failures that precede the k^{th} success.

and noting that $\binom{x-1}{k-1} = \binom{j+k-1}{k-1} = \binom{j+k-1}{j}$,

the series becomes.

$$\sum_{x=k}^{\infty} nb(x;k,p) = \sum_{x=k}^{\infty} \binom{x-1}{k-1} p^k (1-p)^{x-k} = p^k \sum_{j=0}^{\infty} \binom{j+k-1}{j} (1-p)^j$$

$$= p^k \left[1 - (1-p)\right]^{-k} = 1.$$

- Remark: Both $\underline{X}$ = trial number of the k^{th} success

and $\overline{Y}$ = the number of failures that precede the k^{th} Success

$= (\underline{X} - k) \Leftarrow \boxed{\text{text book}}$

are referred, in the literature, to as negative binomial distribution.

- The mgf of $\underline{X} \sim np(k, r, p)$ is

$$M_X(t) = \sum_{x=k}^{\infty} e^{tx} \binom{x-1}{k-1} p^k (1-p)^{x-k}$$

$$= \left[\sum_{j=0}^{\infty} e^{tj} \binom{j+k-1}{j} (1-p)^j\right] e^{tk} p^k$$

$$= e^{tk} p^k \left[1 - e^t (1-p) \right]^{-k}$$

$\underline{\text{Proposition}}$. If $X \sim np(x; k, p)$, then

$$E(X) = \frac{k}{p}, \quad V(X) = \frac{k(1-p)}{p^2}$$

Proof: Differentiating $M_X(t)$ gives

$$M_X^{(1)}(t) = k\, e^{tk} p^k \left[1 - e^t(1-p)\right]^{-k}$$
$$- k\, e^{tk} p^k \left[1 - e^t(1-p)\right]^{-k-1} \left[-e^t(1-p)\right]$$

$$\Rightarrow E(X) = M_X^{(1)}(0) = k + k \cdot p^{-1}(1-p) = k/p$$

$$M_X^{(2)}(t) = k^2 e^{tk} p^k \left[1 - e^t(1-p)\right]^{-k}$$
$$- k^2 e^{tk} p^k \left[1 - e^t(1-p)\right]^{-k-1} \left[-e^t(1-p)\right]$$
$$- k^2 e^{tk} p^k \left[1 - e^t(1-p)\right]^{-k-1} \left[-e^t(1-p)\right]$$
$$+ k(k+1) e^{tk} p^k \left[1 - e^t(1-p)\right]^{-k-2} \left[-e^t(1-p)\right]^2$$
$$- k\, e^{tk} p^k \left[1 - e^t(1-p)\right]^{-k-1} \left[-e^t(1-p)\right]$$

$$\Rightarrow E(X^2) = M_X^{(2)}(0) = k^2 + 2k^2 \cdot p^{-1}(1-p) - k(k+1) \cdot p^{-2}(1-p)^{2}$$
$$+ k \cdot p^{-1}(1-p)$$

$$V(X) = E(X^2) - \left[E(X)\right]^2 \Longrightarrow \frac{k(1-p)}{p^2}$$