Computation of $Var(X)$:

Assume that X is a random variable with probability distribution

X	x_1	$\cdots$	x_N
$p(x)$	$p(x_1)$	$\cdots$	$p(x_N)$

$$
\begin{aligned}
Var(X) &= \sigma^2 = \sum_{i=1}^{N}[x_i - E(X)]^2 p(x_i) \\
&= \sum_{i=1}^{N}[x_i^2 - 2x_iE(X) + (E(X))^2]p(x_i) \\
&= \sum_{i=1}^{N} x_i^2 p(x_i) - 2E(X)\sum_{i=1}^{N} x_i p(x_i) + (E(X))^2 \sum_{i=1}^{N} p(x_i) \\
&= E(X^2) - 2E(X)E(X) + (E(X))^2 \\
&\qquad (\because \sum_{i=1}^{N} x_i p(x_i) = E(X), \quad \sum_{i=1}^{N} p(x_i) = 1) \\
&= E(X^2) - (E(X))^2 = E(X^2) - \mu^2.
\end{aligned}
$$

Mean and variance for Binomial random variable

If $X \sim N(n,p)$, where $N(\cdot,\cdot)$ indicate binomial distribution, then

$$\mu = E(X) = np, \quad \sigma^2 = Var(X) = np(1-p)$$

Proof: Let $X_i = \begin{cases} 1, & p \\ 0, & 1-p \end{cases}$ indicate the i-th Bernoulli trial, then

$$X = \sum_{i=1}^{n} X_i, \quad \text{and}$$

$E(X_i) = 1 \times p + 0 \times (1-p) = p, \quad E(X_i^2) = 1^2 \times p + 0^2 \times (1-p) = p$

$Var(X_i) = E(X_i^2) - (E(X_i))^2 = p - p^2 = p(1-p).$

Since X_i are i.i.d., based on the properties of expectation and variance we have

$$\mu = E(X) = E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} p = np$$

$$\sigma^2 = Var(X) = Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i) = \sum_{i=1}^{n} p(1-p) = np(1-p)$$

Negative Binomial distribution

Duality between Binomial and Negative Binomial:

Binomial distribution:

n — # of Bernoulli trials (fix)

Y — # of successes among n Bernoulli trials (random)

$P(Y = y) = \binom{n}{y}\pi^y(1-\pi)^{n-y}, \;\; y = 0, 1, \cdots, n$

Negative Binomial distribution:

r — # of successes (fix)

Y — # of Bernoulli trials until r successes (random)

$P(Y = y) = \binom{y-1}{r-1}\pi^r(1-\pi)^{y-r}, \;\; y = r, r+1, \cdots$

Geometric distribution: — a special case of Negative Binomial

r =1

Y — # of Bernoulli trials until the first success (random)

$P(Y = y) = \pi(1-\pi)^{y-1}, \;\; y = 1, 2, \cdots$

The Characteristics of Negative Binomial Distribution

Let $Y^* = Y - r$ and $\pi = \frac{r}{\mu+r}$, where $\mu > 0$. Then, Y^* is the number of failures. The probability mass function of Y^* is

$$P(Y^* = y^*) = \binom{y^* + r - 1}{r - 1}\left(\frac{r}{\mu + r}\right)^r\left(1 - \frac{r}{\mu + r}\right)^{y^*}, \; y^* = 0, 1, 2, \cdots .$$

$$E(Y^*) = \mu, \quad Var(Y^*) = \mu + D\mu^2 \geq E(Y^*), \quad (D \geq 0).$$

D is a dispersion parameter. It summarizes the extent of overdispersion relative to the Poisson assumption (variance equals mean).

As $D \to 0$, Negative Binomial $\to$ Poisson.

Overdispersion Phenomenon

Overdispersion — $Var(Y) > E(Y)$ (violate the Posssion assumption)

e.g. Y — # of fatal accidents each week over a year, $E(Y) = 2$.

Due to seasonal fluctuation in intensity, Y displays more variation than that predicted by the Poisson distribution. Hence, the negative binomial distribution should be used to describe the count data in this case.

Multinomial Distribution

Multinomial variable:

e.g. A multinomial variable with $K = 5$ categories:

y	1	2	3	4	5
p	π_1	π_2	π_3	π_4	π_5

Express multinomial outcomes in vector form:

$$
\begin{array}{lcl}
3 - (0,0,1,0,0), & \qquad & 2 - (0,1,0,0,0) \\
5 - (0,0,0,0,1), & & 3 - (0,0,1,0,0) \\
1 - (1,0,0,0,0), & & 4 - (0,0,0,1,0) \\
\vdots & & \vdots
\end{array}
$$

Multinomial distribution:

K — # of categories

n — # of multinomial trials

Y_j — # of "j" that appears in the n multinomial trials (random)

The probability mass function:

$$p(Y_1 = n_1, Y_2 = n_2, \cdots, Y_K = n_K) = \left(\frac{n!}{n_1! n_2! \cdots n_K!}\right) \pi_1^{n_1} \pi_2^{n_2} \cdots \pi_K^{n_K},$$

where

$$\sum_{j=1}^{K} n_j = n, \qquad \sum_{j=1}^{K} \pi_j = 1.$$

A special case of multinomial distribution:

$$Z_j = \begin{cases} 1, & \text{if } Y = j \\ 0, & \text{otherwise,} \end{cases}$$

Z_j	1	0
$P(Z_j)$	π_j	$1 - \pi_j$

Repeat n Bernoulli trials:

$$
\begin{aligned}
P(Z_j = n_j) &= \left(\frac{n!}{n_j!(n-n_j)!}\right) \pi_j^{n_j} (1-\pi_j)^{n-n_j} \\
&= \binom{n}{n_j} \pi_j^{n_j} (1-\pi_j)^{n-n_j}
\end{aligned}
$$