Probability 2 - Notes 6

The Trinomial Distribution

Consider a sequence of n independent trials of an experiment. The binomial distribution arises if each trial can result in 2 outcomes, success or failure, with fixed probability of success p at each trial. If X counts the number of successes, then $X \sim Binomial(n, p)$.

Now suppose that at each trial there are 3 possibilities, say “success”, “failure”, or “neither” of the two, with corresponding probabilities $p,\ \theta,\ 1-p-\theta$, which are the same for all trials. If we write 1 for “success”, 0 for “failure”, and -1 for “neither”, then the outcome of n trials can be described as a sequence of n numbers

$$\omega = (i_1, i_2, ..., i_n),\ \text{where each } i_j \text{ takes vales 1, 0, or -1}$$

Obviously, $P(i_j = 1) = p,\ P(i_j = 0) = \theta\ P(i_j = -1) = 1 - p - \theta$.

Definition. Let X be the number of trials where 1 occurs, and Y be the number of trials where and 0 occurs. The joint distribution of the pare (X,Y) is called the trinomial distribution.

The following statement provides us with .

Theorem. *The joint p.m.f. for (X,Y) is given by*

$$f_{X,Y}(k,l) = P(X = k, Y = l) = \frac{n!}{k!l!(n-k-l)!} p^k \theta^l (1-p-\theta)^{n-k-l},$$

where $k,\ l \geq 0$ and $k + l \leq n$.

Proof. The sample space consists of all sequences of length n described above. If a specific sequence ω has k “successes” (1’s) and l “failures” (0’s)then $P(\omega) = p^k \theta^l (1-p-\theta)^{n-k-l}$. There are $\binom{n}{k}\binom{n-k}{l} = \frac{n!}{k!l!(n-k-l)!}$ different sequences with k “successes” (1’s) and l “failures” (0’s). Hence $P(X = k, Y = l) = \frac{n!}{k!l!(n-k-l)!} p^k \theta^l (1-p-\theta)^{n-k-l}$. $\square$

The name of the distribution comes from the trinomial expansion

$$\begin{aligned} (a+b+c)^n &= (a+(b+c))^n = \sum_{k=0}^{n} \binom{n}{k} a^k (b+c)^{n-k} \\ &= \sum_{k=0}^{n} \sum_{l=0}^{n-k} \binom{n}{k}\binom{n-k}{l} a^k b^l c^{n-k-l} = \sum_{k=0}^{n} \sum_{l=0}^{n-k} \frac{n!}{k!l!(n-k-l)!} a^k b^l c^{n-k-l} \end{aligned}$$

Properties of the trinomial distribution

1) The marginal distributions of X and Y are just $X \sim Binomial(n, p)$ and $Y \sim Binomial(n, \theta)$. This follows the fact that X is the number of “successes” in n independent trials with p being the probability of ‘successes” in each trial. Similar argument works for Y.

Note that therefore $E[X] = np$, $E[Y] = n\theta$ and $E[Y^2] = Var(Y) + (E[Y])^2 = n\theta(1-\theta) + n^2\theta^2$

2) If $Y=l$, then the conditional distribution of $X|(Y=l)$ is $Binomal(n-l, \frac{p}{1-\theta})$.
Proof.

$$\begin{aligned} P(X=k|Y=l) &= \frac{P(X=k,Y=l)}{P(Y=l)} = \frac{\frac{n!}{k!l!(n-k-l)!}p^k\theta^l(1-p-\theta)^{n-k-l}}{\frac{n!}{l!(n-l)!}\theta^l(1-\theta)^{n-l}} \\ &= \binom{n-l}{k}\left(\frac{p}{1-\theta}\right)^k\left(1-\frac{p}{1-\theta}\right)^{n-l-k} \end{aligned}$$

for $x=0,1,...,(n-y)$. Hence $(X|Y=y) \sim Binomial\left(n-y, \frac{p}{1-\theta}\right)$. □

This is intuitively obvious. Consider those trials for which "failure" (or 0) did not occur. There are $(n-l)$ such trials, for each of which the probability that 1 occurs is actually the conditional probability of 1 given that 0 has not occurred, i.e. $\frac{p}{1-\theta}$. So you have the standard binomial set-up.

3) We shall now use the results on conditional distributions (Notes 5) and the above properties to find $Cov(X,Y)$ and the coefficient of correlation $\rho(X,Y)$.

We proved that $E[XY]=E[YE[X|Y]]$ (see the last page of Notes 5). According to property 2), $E[X|Y=l]=(n-l)\frac{p}{1-\theta}$ and thus $E[X|Y]=(n-Y)\frac{p}{1-\theta}$. Hence

$$\begin{aligned} E[XY] &= E\left[Y\times(n-Y)\frac{p}{(1-\theta)}\right] = \frac{p}{1-\theta}E(nY-Y^2) = \frac{p}{1-\theta}\left(n^2\theta-n\theta(1-\theta)-n^2\theta^2\right) \\ &= \frac{p}{(1-\theta)}[n(n-1)\theta(1-\theta)] = n(n-1)p\theta \end{aligned}$$

Therefore $Cov(X,Y)=E[XY]-E[X]E[Y]=n(n-1)p\theta-n^2p\theta=-np\theta$ and hence

$$\rho(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}=\frac{-np\theta}{\sqrt{n^2p(1-p)\theta(1-\theta)}}=-\left(\frac{p\theta}{(1-p)(1-\theta)}\right)^{\frac{1}{2}}$$

Note that if $p+\theta=1$ then $Y=n-X$ and there is an exact linear relation between Y and X. In this case it is easily seen that $\rho(X,Y)=-1$.

Definition of the multinomial distribution

Now suppose that there are k outcomes possible at each of the n independent trials. Denote the outcomes $A_1,A_2,...,A_k$ and the corresponding probabilities $p_1,...,p_k$ where $\sum_{j=1}^k p_j=1$. Let X_j count the number of times A_j occurs. Then

$$P(X_1=x_1,...,X_{k-1}=x_{k-1})=\frac{n!}{x_1!x_2!...x_{k-1}!(n-\sum_{j=1}^{k-1}x_j)!}p_1^{x_1}p_2^{x_2}...p_{k-1}^{x_{k-1}}p_k^{n-\sum_{j=1}^{k-1}x_j}$$

where $x_1,x_2,...,x_{k-1}$ are non-negative integers with $\sum_{j=1}^{k-1}x_j\le n$.