Probability Models - Notes1

Review of common probability distributions

1. Single trial with probability p of success. X counts the number success (so $X = 1$ if the outcome is a success and $X = 0$ if the outcome is a failure). Then one says that $X \sim$ Bernoulli(p). $P(X = 1) = p$ and $P(X = 0) = q$ where $q = 1 - p$. $E[X] = p$, $Var(X) = pq$.

2. Sequence of n independent trials, each with probability p of success. X counts the number of successes. Then $X \sim$ Binomial(n, p). $P(X = k) = \binom{n}{k} p^k q^{n-k}$ for $k = 0, 1, ..., n$. $E[X] = np$, $Var(X) = npq$.

Binomial expansion is $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$. If we let $a = p$ and $b = q$ this shows that $\sum_{k=0}^{n} P(X = k) = (p+q)^n = 1^n = 1$.

3. Sequence of independent trials, each with probability p of success. X counts the number of trials required to obtain the first success. Then $X \sim$Geometric(p). $P(X = k) = q^{k-1}p$ for $k = 1, 2, ...$ $E[X] = \frac{1}{p}$, $Var(X) = \frac{q}{p^2}$.

Sum of geometric series is $\sum_{k=1}^{\infty} az^{k-1} = \frac{a}{(1-z)}$ if $|z| < 1$. If we let $a = p$ and $z = q$, this shows that $\sum_{k=1}^{\infty} P(X = k) = \frac{p}{1-q} = 1$.

4. Sequence of independent trials, each with probability p of success. X counts the number of trials required to obtain the k^{th} success. Then $X \sim$Negative Binomial (k, p). $P(X = n) = \binom{n-1}{k-1} p^k q^{n-k}$ for $x = k, k+1, ...$ $E[X] = \frac{k}{p}$, $Var(X) = \frac{kq}{p^2}$.

For $|z| < 1$ the negative binomial expansion is just

$$(1-z)^{-k} = 1 + kz + \frac{k(k+1)}{2!}z^2 + \frac{k(k+1)(k+2)}{3!}z^3 + ... = \sum_{n=k}^{\infty} \binom{n-1}{k-1} z^{n-k}.$$

Hence if we let $z = q$ then $\sum_{n=k}^{\infty} P(X = n) = p^k(1-q)^{-k} = 1$.

Exercise. Prove that $X = X_1 + X_2 + ... + X_k$, where X_j, $1 \leq j \leq k$ are independent $X \sim$Geometric(p) random variables. Using this fact derive the formulae $E[X] = \frac{k}{p}$, $Var(X) = \frac{kq}{p^2}$.

5. If events occur randomly and independently in time, at rate λ per unit time, and X counts the number of events in a unit time interval then $X \sim$Poisson(λ). $P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$ for $k = 0, 1, ...$ $E[X] = \lambda$, $Var(X) = \lambda$.

Taylor expansion of exponential is $e^z = \sum_{x=0}^{\infty} \frac{z^k}{k!}$. Hence if we let $z = \lambda$ then$\sum_{k=0}^{\infty} P(X = k) = e^{-\lambda}e^{\lambda} = 1$.

Probability Generating Function (p.g.f)

Definition For a discrete random variable X which can only take non-negative integer values we define the probability generating function associated with X to be:

$$G_X(t) = E[t^X]$$

or, equivalently,

$$G_X(t) = \sum_{k=0}^{\infty} P(X=k)t^k, \quad \text{where } |t| \leq 1.$$

Note that this is a power series in t.

We can easily find the p.g.f. for all the common probability distributions 1-5 using the expansions given earlier. Note that the hypergeometric (covered in Probability 1) has no simple form for the p.g.f.

(1) $G_X(t) = q + pt$.

(2) $G_X(t) = \sum_{k=0}^{n} \binom{n}{k} (pt)^k q^{n-k} = (pt+q)^n$.

(3) $G_X(t) = \sum_{k=1}^{\infty} (pt)(qt)^{k-1} = \frac{pt}{1-qt}$.

(4) $G_X(t) = (pt)^k \sum_{n=k}^{\infty} \binom{n-1}{k-1} (qt)^{n-k} = \frac{(pt)^k}{(1-qt)^k}$.

(5) $G_X(t) = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} = e^{\lambda(t-1)}$.

It is easily seen that $G_X(0) = P(X=0)$, $G_X(1) = 1$ and $G_X(t)$ is monotone increasing function of t for $t \geq 0$.

Uses of the p.g.f.

1. Knowing the p.g.f. determines the probability mass function.

The p.g.f., $G_X(t)$, is a power series with the coefficient of t^k just the probability $P(X=k)$. There is a unique power series expansion. Hence if X and Y are two random variables with $G_X(t) = G_Y(t)$, then $P(X=k) = P(Y=k)$ for all $k = 0, 1, \ldots$.

If we know the p.g.f. then we can expand it in a power series and find the individual terms of the probability mass function.

e.g. $G_X(t) = \frac{1}{2}(1+t^2) = \frac{1}{2} + 0 \times t + \frac{1}{2}t^2 + 0 \times t^3 + \ldots$. Hence $P(X=0) = \frac{1}{2}$, $P(X=2) = \frac{1}{2}$ and $P(X=x) = 0$ for all other non-negative integers x.

If we recognise the p.g.f. $G_X(t)$ as a p.g.f. corresponding to a specific distribution, then X has that distribution. We do not need to bother doing the power series expansion!

e.g. if $G_X(t) = e^{2t-2} = e^{2(t-1)}$, this is the p.g.f. for a Poisson distribution with parameter 2. Hence $X \sim$Poisson (2).

2. We can differentiate the p.g.f. to obtain $P(X=r)$ and the factorial moments (and hence the mean and variance of X).

$$P(X=0)=G_X(0);\ P(X=1)=G_X^{'}(0);\ P(X=2)=\frac{1}{2}G_X^{''}(0)$$

In general $P(X=r)=\frac{1}{r!}G_X^{(r)}(0)$ where $G_X^{(r)}(t)=\frac{d^rG_X(t)}{dt^r}$.

Theorem

$$E[X]=G_X^{'}(1);\ \ E[X(X-1)]=G_X^{''}(1);\ \ Var(X)=G_X^{''}(1)+G_X^{'}(1)-[G_X^{'}(1)]^2$$

Proof By definition, $G_X(t)=E[t^X]$. Hence

$$G_X'(t)=\frac{d}{dt}E[t^X]=E[(t^X)']=E[Xt^{X-1}]$$

and therefore $G_X'(1)=E[X]$ which proves the first relation. Similarly,

$$G_X''(t)=\frac{d}{dt}E[Xt^{X-1}]=E[(Xt^{X-1})']=E[X(X-1)t^{X-1}]$$

and therefore $G_X''(1)=E[X(X-1)]$ which proves the second relation. Note that $E[X(X-1)]=E[X^2]-E[X]$ and hence $E[X^2]=E[X(X-1)]+E[X]=G_X^{''}(1)+G_X^{'}(1)$. By the definition of the variance and due to these relations we obtain

$$Var(X)=E[X^2]-(E[X])^2=G_X^{''}(1)+G_X^{'}(1)-[G_X^{'}(1)]^2.\quad\square$$

Remarks. 1. Strictly speaking, the above proof is incomplete. The reason for that is that we interchanged the operation of computing the expectation with that of computing a derivative. The technique which would allow us to justify the above lies beyond the scope of this course.

2. More generally, the r^{th} factorial moment can be computed as follows:

$$E[X(X-1)...(X-r+1)]=G_X^{(r)}(1)$$

3. The other way is to obtain the above relations by differentiating $G_X(t)=P(X=0)+tP(X=1)+t^2P(X=2)+...$ termwise:

$$G_X^{'}(t)=P(X=1)+2tP(X=2)+3t^2P(X=3)+....$$

from which we have $E[X]=G_X^{'}(1)$ and $P(X=1)=G_X^{'}(0)$ and for any positive integer r

$$\frac{d^rG_X(t)}{dt^r}=r!P(X=r)+\frac{(r+1)!}{1!}tP(X=r+1)+\frac{(r+2)!}{2!}t^2P(X=r+2)+...$$

from which we have $E[X(X-1)...(X-r+1)] = G_X^{(r)}(1)$ and $P(X=r) = \frac{G_X^{(r)}(0)}{r!}$ e.g. If $G_X(t) = \frac{1+t}{2}e^{(t-1)}$ find $E[X]$, $Var(X)$, $P(X=0)$ and $P(X=1)$.

$$G_X^{'}(t) = \frac{1}{2}e^{(t-1)} + \frac{1+t}{2}e^{(t-1)}$$

$$G_X^{(2)}(t) = \frac{1}{2}e^{(t-1)} + \frac{1}{2}e^{(t-1)} + \frac{1+t}{2}e^{(t-1)}$$

Hence $E[X] = G_X^{'}(1) = \frac{3}{2}$, $var(X) = G_X^{(2)}(1) + \frac{3}{2} - \frac{9}{4} = \frac{5}{4}$, $P(X=0) = G_X(0) = \frac{e^{-1}}{2}$ and $P(X=1) = G_X^{'}(0) = e^{-1}$.

3. Using the p.g.f. to find the distribution of the sum of two or more independent random variables.

Recall that if X and Y are independent random variables then $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$.

Theorem. *Let X and Y be independent random variables with p.g.f.'s $G_X(t)$ and $G_Y(t)$. Then $Z = X+Y$ has p.g.f. $G_Z(t) = G_X(t)G_Y(t)$.*

Proof. Using the independence we can write

$$G_Z(t) = E[t^Z] = E[t^{X+Y}] = E[t^X t^Y] = E[t^X]E[t^Y] = G_X(t)G_Y(t).$$

□

This extends to the sum of a fixed number n of independent random variables.

If $X_1, ..., X_n$ are independent and $Z = \sum_{j=1}^{n} X_j$ then

$$G_Z(t) = \prod_{j=1}^{n} G_{X_j}(t).$$

Example 1. Let X and Y be independent random variables with $X \sim Binomial(n,p)$ and $Y \sim Binomial(m,p)$ and let $Z = X+Y$. Then

$$G_Z(t) = G_X(t)G_Y(t) = (pt+q)^n(pt+q)^m = (pt+q)^{m+n}$$

This is the p.g.f. of a binomial random variable. Hence $Z \sim Binomial(n+m,p)$.

Example 2. Let $X_1, ..., X_m$ be m independent random variables with $X_j \sim Binomial(n_j,p)$ and let $Z = \sum_{j=1}^{m} X_j$ and $N = \sum_{j=1}^{m} n_j$. Then

$$G_Z(t) = \prod_{j=1}^{m} G_{X_j}(t) = \prod_{j=1}^{m} (pt+q)^{n_j} = (pt+q)^N$$

This is the p.g.f. of a binomial random variable. Hence $Z \sim Binomial(N,p)$.