Economics 520, Fall 2005

Lecture Note 5: Special Distributions (CB 3.1-3.3)

Let us look at a number of distributions that are of special importance. Often these are families of distributions, where the PDF or PMF depends on a parameter or parameters that can take on a number of different values. This increases the flexibility of the distribution to fit particular applications. For example the binomial distribution has parameters n and p, so it can be applied to cases with different numbers of trials and with different probabilities of success. We divide the special distributions into discrete and continuous distributions.

Discrete Distributions

1. Binomial Distribution The PMF for the binomial distribution with parameters n and p, denoted by $\mathcal{B}(n,p)$, is

$$f_X(x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{(n-x)},$$

for $x = 0, 1, 2, \ldots, n$, and zero otherwise. We have already seen that the mean of X is np. The variance can be calculated in the same way. Alternatively, one can think of the interpretation of X as the sum of repeated Bernoulli trials. A single Bernoulli trial has a binomial $\mathcal{B}(1,p)$ distribution with PMF

$$f_Y(y) = p^y(1-p)^{1-y},$$

for $y = 0, 1$. Its mean is p and its variance is

$$V(Y) = E(Y^2) - (E(Y))^2 = E(Y) - E(Y) \cdot E(Y) = p(1-p).$$

The variance of a sum of independent random variables is equal to the sum of the variances. So, the variance of $X = \sum_{i=1}^n Y_i$ is

$$V(X) = V\left(\sum_{i=1}^n Y_i\right) = \sum_{i=1}^n V(Y_i) = np(1-p).$$

2. Poisson Distribution The PMF of the Poisson distribution with parameter $\lambda > 0$, is

$$f_X(x) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!} & x = 0, 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

(Note that unlike the binomial distribution, there is no explicit upper bound to X.) The MGF is

$$M_X(t) = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x e^{tx}}{x!} = \sum_{x=0}^{\infty} \frac{e^{-\lambda}(\lambda e^t)^x}{x!}$$

$$= \exp(e^t\lambda - \lambda) \cdot \sum_{x=0}^{\infty} \frac{e^{-e^t\lambda}(\lambda e^t)^x}{x!} = \exp(e^t \cdot \lambda - \lambda).$$

The cumulant generating function is

$$K_X(t) = \ln M_X(t) = e^t \cdot \lambda - \lambda.$$

All the cumulants are equal to λ, and hence the mean and variance of X are

$$K_X^1(0) = \lambda, \qquad \text{and } K_X^2(0) = \lambda.$$

There is a close connection between the binomial distribution and the Poisson distribution. Suppose we observe for a period of time of length 1 a system in which events of a particular type occur. For example, we may observe a telephone exchange with calls coming in, or a radioactive substance emitting radioactive particles. It may be reasonable to assume that the probability of an event occurring in a short period of time Δ is roughly proportional to the length of time, so

$$Pr\Big(\text{event in } (t, t+\Delta)\Big) = \lambda\Delta + o(\Delta).$$

(The $o(\Delta)$ notation implies that the remainder is small in the sense that $\lim_{\Delta\to 0} o(\Delta)/\Delta = 0$.) Similarly, it may be reasonable to assume that the probability of two events occurring in a short period of time is small, or

$$Pr\Big(\text{two events in } (t, t+\Delta)\Big) = o(\Delta).$$

Finally, we assume that the events occuring in different periods of time are independent:

$$Pr\Big(\text{event in } (t,s) \underline{\text{and}} \text{ event in } (u,v)\Big) = Pr\Big(\text{event in } (t,s)\Big) \times Pr\Big(\text{event in } (u,v)\Big),$$

for s, t, u, v such that $(u,v) \cap (s,t) = \emptyset$. We could try to model the number of events in the period from 0 to 1 through a binomial distribution by chopping up the interval into n intervals of length $\Delta = 1/n$. In each interval the probability of the event occurring is approximately $p \approx \lambda\Delta = \lambda/n$. Then the distribution of X, the total number of periods in which an event occurred, is

$$f_X(x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{(n-x)} = \frac{n!}{x!(n-x)!} \cdot (\lambda/n)^x \cdot (1-\lambda/n)^{(n-x)},$$

for $x = 0, 1, 2, \ldots, n$, and zero otherwise. As n goes to infinity, the number of periods in which an event occurs, X, approaches the number of events occurring in the interval from 0 to 1. Take the limit as n goes to infinity, for fixed x:

$$\lim_{n\to\infty} \frac{n!}{(n-x)! \cdot n^x} = \lim_{n\to\infty} \prod_{y=0}^{x-1} \frac{n-y}{n} = \prod_{y=0}^{x-1} \lim_{n\to\infty} \frac{n-y}{n} = 1^x = 1,$$

$$\lim_{n\to\infty} (1-\lambda/n)^{-x} = 1.$$

Also, recall that $\lim_{n\to\infty} (1+a/n)^n = e^a$, so

$$\lim_{n\to\infty} (1-\lambda/n)^n = e^{-\lambda}.$$

Hence

$$\lim_{n\to\infty} \frac{n!}{x!(n-x)!} \cdot (\lambda/n)^x \cdot (1-\lambda/n)^{(n-x)} = \frac{e^{-\lambda}\lambda^x}{x!},$$

which is the PMF for the Poisson distribution. Hence the Poisson distribution is a good approximation to a binomial distribution with large n and small p. It also explains why λ is referred to as the arrival rate.

3. Geometric Distribution Given the same type of experiment with independent Bernoulli trials, each with probability of success p as in the binomial distribution, we can ask a different question. How many trials does it take to the first success? The probability of the first success occurring on the first trial is p, on the second $p(1-p)$ etcetera. Hence the pmf for a random variable X with a geometric distribution is

$$f_X(x) = p \cdot (1-p)^{x-1},$$

for $x = 1, 2, \ldots$ and zero elsewhere. The mean and variance are $1/p$ and $(1-p)/p^2$ respectively.

4. Negative Binomial Distribution An extension of the geometric distribution considers the distribution of the number of trials required for the rth success. The only scenario for this is to have a success on the xth trial, which has probability p, and $r-1$ successes in the first $x-1$ trials and $x-r$ failures in the first $x-1$ trials The latter has probability

$$\binom{x-1}{r-1} \cdot p^{r-1} \cdot (1-p)^{(x-r)}.$$

Hence the pmf for the negative binomial distribution is

$$f_X(x) = \binom{x-1}{r-1} \cdot p^r \cdot (1-p)^{(x-r)},$$

for $x = r, r+1, r+2, \ldots$, and zero otherwise. The mean and variance for X are $r + r(1-p)/p$ and $r(1-p)/p^2$ respectively. Often instead of X the focus is on $Y = X - r$, the number of trials in excess of r. Its mean and variance are $E(Y) = E(X) - r = r(1-p)/p$ and $V(Y) = V(X) = r(1-p)/p^2$ respectively.

Continuous Distributions

1. Uniform Distribution A random variable X has a uniform distribution on the interval (a, b) if its PDF is

$$f_X(x) = \frac{1}{b-a}, \qquad \text{for } x \in (a, b), \text{ and 0 otherwise.}$$

The mean and variance are $(a+b)/2$ and $(b-a)^2/12$ respectively.

2. Exponential Distribution The PDF for a random variable X with an exponential distribution with parameter $\lambda > 0$ is

$$f_X(x) = \lambda \exp(-\lambda x) \qquad \text{for } x > 0 \text{ and 0 otherwise.}$$

Alternatively, the exponential distribution is written in terms of a parameter $\beta = 1/\lambda$, so the PDF is

$$f_X(x) = (1/\beta) \exp(-x/\beta) \qquad \text{for } x > 0 \text{ and 0 otherwise.}^1$$

[1] In previous lecture notes I used θ instead of λ. Switching to λ is convenient to see how the exponential distribution is related to the Poisson.

The moment generating function for this distribution is

$$M_X(t) = \int_0^\infty \lambda \exp(-\lambda x + tx) = \int_0^\infty \lambda \exp(-x(\lambda - t))$$

$$= \frac{\lambda}{\lambda - t} \int_0^\infty (\lambda - t) \exp(-x(\lambda - t)) = \frac{\lambda}{\lambda - t},$$

for $|t| < \lambda$. The cumulant generating function is

$$K_X(t) = \ln M_X(t) = \ln \lambda - \ln(\lambda - t),$$

with

$$K_X^1(t) = \frac{1}{\lambda - t}, \quad \text{and} \quad K_X^2(t) = \frac{1}{(\lambda - t)^2}.$$

The mean and variance follow easily from this as $1/\lambda$ and $1/\lambda^2$.

An important connection exists with the same example that motivated the relation between the Poisson and binomial distributions. In this example of observing a system where events occurr at random times, define the random variable X as the waiting time until the first event. With the probability of the event occurring in a period of length $\Delta = 1/n$ approximately equal to $p \approx \lambda\Delta = \lambda/n$, the probability of no event occurring till time t (that is, in the first $t/\Delta = tn$ periods), is $(1 - \lambda/n)^{tn}$. Take the limit as n goes to infinity to get $\exp(-\lambda t)$. Hence the CDF for the distribution of the waiting time is $F_X(x) = 1 - \exp(-\lambda x)$ and the PDF is $\lambda \exp(-\lambda x)$, which is the PDF for an exponential distribution.

An important property of the exponential distribution, and one that is intuively obvious from the above motivation is that of lack of memory. The probability of X exceeding c is $\exp(-\lambda c)$. Now consider conditioning on the event that we have already waited a period of length d. Then the probability of having to wait for another period of length c is

$$Pr(X > c + d | X > d) = \frac{Pr(X > c + d, X > d)}{Pr(X > d)} = \frac{Pr(X > c + d)}{Pr(X > d)}$$

$$= \frac{\exp(-\lambda(c + d))}{\exp(-\lambda d)} = \exp(-\lambda c),$$

and similarly

$$E[X - d | X > d] = 1/\lambda,$$

for all $d \geq 0$. Having waited already for a period of length d does not change how long one is expected to have left to wait. There is no information in the fact that we have already waited for a long period of time. In some applications, this memoryless property is sensible. For example waiting for the next time the roulette wheel gives a 37 is not affected by how long we have waited already (if it is a fair roulette wheel!). However, for other types of phenomenon where the memoryless property is not appropriate, the exponential distribution may not be an appropriate probability model.