Classroom Notes

Math 394B Summer 2005

Week 6

1 Some Consideration On Problems in Assignment Due 8/1

1.1 Using Cumulative Distribution Functions To Compute Probabilities

We know that the definition of cdf of a variable X is $F_X(x) = P[X \leq x]$. It follows by the usual basic rules that

$$P[a < X \leq b] = F_X(b) - F_X(a)$$

But we usually want the probability of a closed or open interval - not a semi-open one! To solve this "problem" we use the axiom of *continuity* for probabilities, and the fact that

$$\{a < x < b\} = \bigcup_{n=1}^{\infty} \left\{a < x \leq b - \frac{1}{n}\right\}$$

$$\{a \leq x \leq b\} = \bigcap_{n=1}^{\infty} \left\{a - \frac{1}{n} < x \leq b - \frac{1}{n}\right\}$$

(you may recall from set theory that the *intersection* of an arbitrary number of closed intervals is closed, but not their union, while the *union* of an arbitrary number of open intervals is open, but not their intersection).

Using these facts, we find that

$$P[a < X < b] = \lim_{n\to\infty} F_X\left(b - \frac{1}{n}\right) - F_X(a)$$

$$P[a \leq X \leq b] = F_X(b) - \lim_{n\to\infty} F_X\left(a - \frac{1}{n}\right)$$

(we wrote $\frac{1}{n}$ for a quantity that goes to zero taking a countable number of values - that was a technical fussiness. It is true that $b - \frac{1}{n}$ might be smaller than a at first, and the expression would be meaningless, for the first few values of n: in that case let the limiting sequence start with $n$big enough...)

1.2 Interesting Facts About Variances

In a problem we prove, on the side, the following interesting fact: for all a and b

$$Var\left[aX+b\right]=a^{2}Var\left[X\right]$$

In class, we also proved another interesting fact: for any two RV's

$$Var\left[X+Y\right]=Var\left[X\right]+Var\left[Y\right]+2Cov\left[X,Y\right]$$

where the *covariance* of two RV's is defined as

$$Cov\left[X,Y\right]=E\left[\left(X-EX\right)\left(Y-EY\right)\right]$$

Let us observe that, at least for discrete RV's,

$$E\left[\left(X-EX\right)\left(Y-EY\right)\right]=\sum_{x,y}\left(x-EX\right)\left(y-EY\right)P\left[X=x,Y=y\right] \qquad (1)$$

i.e., we need to know the *joint* distribution of X and Y, which enters crucially in the formula.

However, if the covariance happened to be zero, we would have the attractive expression "the variance of the sum is the sum of the variances". If two variables have zero covariance, they are said to be *uncorrelated.* Most variables *are* correlated, but it is always nice to spot situations when they are not. A notable case is the following:

Proposition: Independent RV's are uncorrelated.

Proof: Just observe that independence means $P\left[X=x,Y=y\right]=P\left[X=x\right]P\left[Y=y\right]$, and work in a straightforward way on (1).

In general, the converse is *completely false.* A trivial example is the following: let X be such that $EX=0=E\left[X^{3}\right]$ (it is easy to write as many examples as you wish of such distributions). Now consider X and $Y=X^{2}$. Since $Y=f\left(X\right)$, it is pretty clear that they are *not* independent. But $EX=0$, and

$$E\left[\left(X-EX\right)\left(Y-EY\right)\right]=E\left[X\left(X^{2}-E\left[X^{2}\right]\right)\right]=E\left[X^{3}-X\cdot E\left[X^{2}\right]\right]=$$

$$=E\left[X^{3}\right]-EX\cdot E\left[X^{2}\right]=0$$

Especially in engineering, you will often find that people tend to treat uncorrelated variables as if they were independent. In general they are not, as we just observed. It is however true that, in very special but important cases, uncorrelated variables *are* indeed independent. The most important of these special cases is when the variables have a *joint Gaussian distribution* (hopefully, we'll explain what this means in the next few days).

1.3 The Poisson Distribution As a Limit

The argument that shows that a binomial variable with large n and small p can be approximately described by a Poisson distribution with parameter np, is sometimes called "The Law of Rare Events": rare events over a large number of experiments look Poisson. This is why, for instance, radioactive decay is usually assumed to obey a Poisson distribution.

One other way of proving this limit law is to take for granted the following (non trivial) theorem, which refers to the *Moment Generating Function,* the *Characteristic Function,* and the *Generating Function* introduced in the notes for week 2.

Theorem Let M be one of the aforementioned functions. Suppose we have a sequence of RVs X_n, with corresponding M_n, and that $M_n \to M$, pointwise, with M being the corresponding function of a distribution. Let X be a RV with this distribution. Then

$$P\left[a \le X_n \le b\right] \to P\left[a \le X \le b\right]$$

for any a, b.

Note that it is the *probabilities* that converge. The RVs, seen as functions on Ω most likely do not converge in any sense at all.

We already computed the MGF, the CF, and the GF of binomial variables. Here we do the same for a Poisson variable, after which it is easy to reprove the Law or Rare Events. Suppose X is Poisson, with parameter λ:

1. $Ee^{tX} = \sum_{k=0}^{\infty} e^{tk} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\left(\lambda e^t\right)^k}{k!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda\left(e^t - 1\right)}$
2. $Ee^{itX} = \sum_{k=0}^{\infty} e^{itk} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\left(\lambda e^{it}\right)^k}{k!} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda\left(e^{it} - 1\right)}$
3. $Ez^X = \sum_{k=0}^{\infty} \left(\frac{\lambda^k}{k!} e^{-\lambda}\right) z^k = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda z)^k}{k!} = e^{-\lambda} e^{\lambda z} = e^{\lambda(z-1)}$

1.4 Moments of A Poisson RV

We can use the results in sec. 1.3 to compute moments of a Poisson RV. For instance, expanding the MGF in powers of t we find

$$e^{\lambda\left(e^t-1\right)} = 1 + \sum_{k=1}^{\infty} \frac{\lambda^k \left(e^t - 1\right)^k}{k!} = 1 + \lambda\left(e^t - 1\right) + \frac{\lambda^2}{2}\left(e^t - 1\right)^2 + \ldots =$$

$$1 + \lambda\left(t + \frac{t^2}{2} + \ldots\right) + \frac{\lambda^2}{2}\left(e^{2t} - 2e^t + 1\right) = 1 + \lambda\left(t + \frac{t^2}{2} + \ldots\right) + \frac{\lambda^2}{2}\left(2t + \frac{4t^2}{2} - 2t - \frac{2t^2}{2} + \ldots\right) =$$

$$= 1 + \lambda t + \frac{t^2}{2}\left(\lambda + \lambda^2\right) + \ldots$$

and, looking back at the expansion of the MGF, we see that $EX = \lambda$, which we knew by other means, and $EX^2 = \lambda + \lambda^2$. Of course, we could have found the same result by computing $\frac{dM(0)}{dt}$, and $\frac{d^2M(0)}{dt^2}$. Anyway, we end up with $Var[X] = EX^2 - (EX)^2 = \lambda$ (as noted, in a non rigorous way, in class). Pushing the calculations to the 3rd power we could get the third moments, and so on.

2 Clever Tricks To Calculate Moments

We checked in class what the expected value and the variance are for the binomial and Poisson distribution by some quick tricks - we used the fact that the sum of independent Bernoulli RVs has a binomial distribution, and we used the "Law of Rare Events" as a justification for the moments of the Poisson distribution. However, with some effort, we can find these results directly. The following calculations are presented as an illustration of the type of tricks that we can use to compute non-obvious sums.

2.1 The Binomial Distribution

To find the expected value form the definition, we have to compute

$$EX = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} k \binom{n}{k} \left(\frac{p}{1-p}\right)^k (1-p)^n =$$

$$= \frac{p}{1-p}(1-p)^n \sum_{k=0}^{n} k \binom{n}{k} \left(\frac{p}{1-p}\right)^{k-1}$$

Now, if we set $\frac{p}{1-p} = u$, the sum reads like[1]

$$\sum_{k=0}^{n} \binom{n}{k} \frac{du^k}{du} = \frac{d}{du} \sum_{k=0}^{n} \binom{n}{k} u^k = \frac{d}{du}(1+u)^n = n(1+u)^{n-1}$$

and we have

$$EX = p(1-p)^{n-1} n \left(1 + \frac{p}{1-p}\right)^{n-1} = np \frac{(1-p)^{n-1}(1-p+p)^{n-1}}{(1-p)^{n-1}} =$$

$$= np$$

But we already knew that...

[1] We are using Newton's binomial formula: $\sum_{k=0}^{n} \binom{n}{k} u^k$ is the same as $\sum_{k=0}^{n} \binom{n}{k} u^k 1^{n-k} = (1+u)^n$.

One way to compute EX^2 in the same spirit, is to write $k^2 = k(k-1) + k$ so that

$$EX^2 = \sum_{k=0}^{n} k^2 \left(\begin{array}{c} n \\ k \end{array} \right) p^k (1-p)^{n-k} =$$

$$= \sum_{k=0}^{n} k(k-1) \left(\begin{array}{c} n \\ k \end{array} \right) p^k (1-p)^{n-k} + \sum_{k=0}^{n} k \left(\begin{array}{c} n \\ k \end{array} \right) p^k (1-p)^{n-k}$$

The second sum is $EX = np$. The first can be computed using the same trick as before - the term k simply suggests we move to the second derivative...

$$\sum_{k=0}^{n} k(k-1) \left(\begin{array}{c} n \\ k \end{array} \right) p^k (1-p)^{n-k} = \left(\frac{p}{1-p} \right)^2 (1-p)^n \sum_{k=0}^{n} k(k-1) \left(\begin{array}{c} n \\ k \end{array} \right) \left(\frac{p}{1-p} \right)^{k-2}$$

and, setting $\frac{p}{1-p} = u$,

$$\sum_{k=0}^{n} k(k-1) \left(\begin{array}{c} n \\ k \end{array} \right) \left(\frac{p}{1-p} \right)^{k-2} = \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) \frac{d^2 u^k}{du^2} = \frac{d^2}{du^2} \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) u^k =$$

$$= \frac{d^2}{du^2} (1+u)^n = n(n-1)(1+u)^{n-2} = n(n-1) \left(1 + \frac{p}{1-p} \right)^{n-2}$$

Finally,

$$EX^2 = \left(\frac{p}{1-p} \right)^2 (1-p)^n n(n-1) \left(\frac{1-p+p}{1-p} \right)^{n-2} + np =$$

$$n(n-1)p^2 + np$$

and

$$Var[X] = EX^2 - (EX)^2 = n(n-1)p^2 + np - n^2p^2 = -np^2 + np = np(1-p)$$

2.2 The Geometric Distribution

In the same spirit, we can easily compute expected value and variance of the geometric distribution.

$$EX = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = p \sum_{k=1}^{\infty} \left(-\frac{d}{dp} (1-p)^k \right) = -p \frac{d}{dp} \left(\sum (1-p)^k \right) =$$

$$= -p \frac{d}{dp} \left(\frac{1}{1-(1-p)} \right) = -p \left(-\frac{1}{p^2} \right) = \frac{1}{p}$$

As discussed in class, we used the formula

$$1 + q + q^2 + \ldots + q^k = \frac{1-q^{k+1}}{1-q}$$

and $\lim_{n \to \infty} q^n = 0$ when $0 < q < 1$.

As you can imagine, to compute the variance (and get the result $\frac{1-p}{p^2}$), we repeat the very same path we followed in sec. 2.1, including the convenient $k^2 = k(k-1) + k$.

2.3 The Poisson Distribution

Let's conclude with the easy calculations referring to the Poisson distribution:

$$EX = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} =$$

$$\lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

$$EX^2 = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} =$$

$$= \sum_{k=0}^{\infty} k\,(k-1)\,\frac{\lambda^k}{k!} e^{-\lambda} + \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} =$$

$$\lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} + \lambda =$$

$$\lambda^2 e^{-\lambda} e^{\lambda} + \lambda = \lambda^2 + \lambda$$

and the variance is

$$\lambda^2 + \lambda - \lambda^2 = \lambda$$

As a side remark, we changed the lower limits in the sums, by observing that a sum starting at $k = 0$ which includes a factor of k might as well start at 1, since the first term is zero. Same argument for a sum starting at $k = 0$ with factors $k(k-1)$: we might as well start it at 2, since the first two terms are zero. Obviously, finally,

$$\sum_{k=2}^{\infty} A(k-2) = \sum_{j=0}^{\infty} A(j)$$

after setting $j = k - 2$, which explains the result of the sums in the preceding calculations.