Stats 150, Fall 2019
Lecture 8, Tuesday, 9/24/2019

1 §1.11 Infinite State Space

This is starred in the text but is not optional for our course. We will discuss techniques for both finite and infinite state spaces, especially
- probability generating functions
- potential kernel (AKA) Green matrix

2 Review of Math Background

Know the following by heart (we'll need to use them on the midterm).

2.1 Binomial Theorem

This is to write:

$$(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k.$$

We should observe that $\binom{n}{k} = \frac{n!}{(n-k)!k!} = \boxed{\frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1}}$, and it is important that the numerator is a polynomial in n. Pitman comments that no one realized why this is important until about 1670. The reason is that this form can be extended to other powers, namely $n := -1, \frac{1}{2}, \frac{-1}{2}$, or any real number $n \to r \in \mathbb{R}$.
Take r real and look at

$$(1+x)^r = \sum_{k=0}^{\infty} \binom{r}{k} x^k,$$

if $|x| < 1$ for real or complex x. Notice that the conventional r 'choose' k makes sense particularly through the polynomial definition of the binomial factor $\binom{r}{k}$.
This is the instance with $f(x) = x^r$. Now if we want to consider:

$$f(1+x) = f(1) + f'(1)x + \frac{f''(1)}{2!}x^2 + \cdots,$$

which is our familiar Taylor expansion, for $|x| < R$, where R is our radius of convergence. Usually for our purposes, $R \geq 1$.
Now of course, recall

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

We get exponentials as limit of binomial probabilities (e.g. the Poisson distribution). Also, recall that the geometric distribution converges to the exponential distribution with suitable scales.

3 Probability Generating Functions

Suppose we have a nonnegative integer-valued random variable X, which for simplicity will have nonnegative integer values $X \in \{0, 1, 2, \dots\}$.
We define the PGF (probability generating function) of X to be the function

$$\boxed{\phi_X(z) := \mathbb{E}z^X}$$

Pitman says that we usually take $0 \leq z \leq 1$. When discussing PGFs, we may push z beyond this but we will keep it within this bound. Now we try to write the above in terms of a power series. Recall that $\mathbb{E}g(X) = \sum_{n=0}^{\infty} \mathbb{P}(X = n)g(X)$, so

$$\phi_X(z) := \mathbb{E}Z^X = \sum_{n=0}^{\infty} \mathbb{P}(X = n)z^n = \sum_{n=0}^{\infty} P_n z^n.$$

We worked with PGFs very briefly in a previous lecture, namely taking X uniform on $\{1, 2, 3, 4, 5, 6\}$, and we looked at:

$$\phi_X(z) = \frac{1}{6}\left(z + \cdots + z^6\right).$$

Recall this is where Pitman asked us to look this expansion up in Wolfram Alpha.
Notice that by convention, $0^0 = 1$, so $\phi_X(0) = \mathbb{P}(X = 0)$.
Now for a poisson PGF, we have:

$$\begin{aligned}
\frac{d}{dx}\phi_X(z) &= \frac{d}{dz}\sum_n \mathbb{P}(X = n)z^n \\
&= \sum_n \mathbb{P}(X = n)\frac{d}{dz}z^n \\
&= \sum_n \mathbb{P}(X = n)nz^{n-1},
\end{aligned}$$

and so we see that

$$\mathbb{E}X = \frac{d}{dz}\phi_X(3)|_{z=1^-},$$

where we approach from the left if we need to be pedantic.
Perhaps we'd like to compute the variance. We ask, what happens if we differentiate twice?

$$\left(\frac{d}{dz}\right)^2 \phi_x(z) = \sum_{n=0}^{\infty} \mathbb{P}(X = n)n(n-1)z^{n-2}.$$

Again we'd like the z factor to go away, so we set $z := 1$ and we have:

$$\begin{aligned}
\mathbb{E}\left[X(X-1)\right] &= \sum_{n=0}^{\infty} \mathbb{P}(X = n)n(n-1) \\
&= \left(\frac{d}{dz}\right)^2 \phi_X(z)|_{z=1^-}
\end{aligned}$$

Recall that $X_\lambda \sim \text{Poisson}(\lambda)$ if and only if:

$$\mathbb{P}(X_\lambda = n) = \frac{e^{-\lambda}\lambda^n}{n!},$$

which via the generating function implies:

$$\phi_{X_\lambda}(s) = \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n s^n}{n!} = e^{-\lambda}e^{\lambda s} = e^{\lambda(s-1)}.$$

Now we go back to our above to derive (or simply recall):

$$\begin{aligned} \mathbb{E}X_\lambda &= \lambda \\ \mathrm{Var}(X_\lambda) &= \lambda \end{aligned}$$

A (good) question arises whether $\phi_X(z)$ is a probability. The answer is yes, because after all the range of values is between 0 and 1, and any such function can be interpreted as a probability. Notably, we have:

$$\phi_X(z) = \mathbb{P}(X \leq G_{1-z}),$$

where G denotes the geometric density function. Then

$$\mathbb{P}(G_p = n) = (1-p)^n p, \text{ and } \mathbb{P}(G_p \geq n) = (1-p)^n.$$

In summary, we can think of a probability generating function as a probability, and we only need that G_{1-z} is idependent of X.
Now if X, Y are independent, then

$$\begin{aligned} \mathbb{E}z^{X+Y} &= \mathbb{E}\left[z^X z^Y\right] \\ &= \left[\mathbb{E}z^X\right]\left[\mathbb{E}z^Y\right] \\ &= \phi_X(z)\phi_Y(z). \end{aligned}$$

Hence the PGF of a sum of independent variables is the product of their PGFs.

Example: Let $G_p \sim$ Geometric(p) on $\{0, 1, 2, \dots\}$. Then

$$\mathbb{P}(G_p = n) = (1-p)^n p, \text{ for } n = 0, 1, 2, \dots$$

Now if we want to look at the probability generating function, we have:

$$\mathbb{E}(z^{G_p}) = \sum_{n=0}^{\infty} q^n p z^n = \frac{p}{1-qz},$$

for $p + q = 1$ and $|z| < 1$. Now we look at $T_r := G_1 + G_2 + \cdots + G_r$, where $r = 1, 2, 3, \dots$, and G_i are all independent geometrically distributed with the same parameter p.
The interpretation is to see G_p as the waiting time (of the number of failures) before the first success. That is, the number of 0s before the first 1 in independent Bernoulli(p) 0/1 trials. Then similarly,

$$T_r = T_{r,p} = \text{ number of 0s before } r\text{th 1 in indep. Bernoulli(p) 0/1 trials.}$$

Looking at iid copies of G_p we use generating functions:

$$\begin{aligned} \mathbb{E}z^{T_r} &= \left(\frac{p}{1-qz}\right)^r = p^r(1-qz)^{-r} \\ &= p^r\left(1 + (-qz)\right)^{-r} \\ &= \sum_{n=0}^{\infty} \binom{-r}{n}(-qz)^n, \end{aligned}$$

where we simply plug into Newton's binomial formula. Notice that this actually is equal to:

$$\mathbb{E}z^{T_r} = p^r \sum_{n=0}^{\infty} \frac{(r)_{n\uparrow}}{n!} q^n z^n,$$

where we define:

$$\begin{aligned}(r)_{n\uparrow} &= r(r+1)\cdots(r+n-1) \\ \frac{(r)_{n!}}{n!} &= \binom{r+n-1}{n}.\end{aligned}$$

From 134, we know this to be the negative binomial distribution.