Handout 13: Stochastic Processes

Bernoulli Process

A stochastic process is nothing more than a fancy way of describing a collection of random variables. Often this collection is indexed by the variable t and the collection is visualized as being index over time. Many of the distributions we have explored can be defined in terms of stochastic processes, and in many cases we have already more-or-less show this. For example, consider an infinite sequence of independent Bernoulli trials:[1]

[1] This is what is meant when people refer to a Bernoulli process.

$$X_1, X_2, \ldots \sim_{i.i.d.} Bernoulli(p)$$

The main goal of today is to show how every named distribution we have seen can be related to this specific stochastic process. The binomial distribution can be defined as:

$$B = \sum_{t=1}^{n} X_t \sim Bin(n, p)$$

And the Geometric is given by:

$$G = \arg\max_t \{t \text{ such that } X_s = 0, \;\forall s \leq t\} \sim Geom(p)$$

Finally, for completeness, take the following process:

$$G_1, G_2, \ldots \sim_{i.i.d.} Geom(p)$$

The negative binomial distribution can be defined as:

$$N = \sum_{t=1}^{k} G_t \sim NB(k, p)$$

Using the theory of moment generating functions, we can use these stochastic process definitions to quickly calculate the moments of these distributions that are all ultimately related to the Bernoulli distribution.

We can also consider the limiting distribution of a stochastic process. For example, consider the stochastic process:

$$X_1, X_2, \ldots \text{ where } X_t \sim Bin(t, \lambda/t), \quad \text{for a fixed } \lambda \in (0, \infty)$$

We have already seen that in this case the process limits to a Poisson distribution:

$$X_t \to X \sim Poisson(\lambda)$$

The mean of a geometric distribution is $1/p$, so let's define a variable:

$$X_t = Y_t/t, \quad Y_t \sim Geom(\lambda/t)$$

The mean of X_t will then be λ^{-1} for any t. It turns out that the following limit is equal to the exponential distribution:[2]

[2] A derivation of this will be on today's worksheet.

$$X_t \to X \sim Exp(\lambda).$$

For this reason, the exponential distribution is used to model the inter-arrival time of a process modeled by a Poisson distribution.[3] Completing our story, the Gamma distribution comes from taking the following stochastic process:

[3] Such as the time between patients arriving at an ER center.

$$E_1, E_2, \ldots \sim_{i.i.d.} Exp(\lambda)$$

Then:

$$G = \sum_{t=1}^{k} E_t \sim Gamma(k, \lambda^{-1})$$

This yields the following table of relationships:

	Num. of Events	Time Between Events	Time Between k Events
fixed n	$Bin(n,p)$	$Geom(p)$	$NB(k,p)$
fixed $\lambda = np$	$Poisson(\lambda)$	$Exp(\lambda)$	$Gamma(k, \lambda^{-1})$

In the second row the limit is taken as n goes to infinity.

Normal and Beta Distributions

The two distributions we have not shown the relationship to a Bernoulli process yet are the Normal and Beta distributions. We actually already saw the relationship for the Beta distribution on the previous worksheet, but never gave this relationship a name. As a reminder, if:

$$p \sim Beta(\alpha, \beta)$$
$$X|p \sim Bin(n,p)$$

Then:

$$p|X \sim Beta(k + \alpha, n - x + \beta)$$

Because the family of the distribution of p is the same whether or not I condition on X, we say that the Beta and Binomial (or Bernoulli, as a special case) are *conjugate priors*. This is a concept we will return to at length following the second exam.

Now, what about the Normal distribution? We can get the normal by considering a Bernoulli process:

$$X_1, X_2, \ldots \sim_{i.i.d.} Bernoulli(p)$$

And defining:

$$Z_t = \frac{\frac{1}{n} \cdot \sum_i (X_i - p)}{p(1-p)}$$

The variable Z_t should have zero mean and a variance of 1 for every value of t. Then:

$$Z_t \to N(0, 1).$$

We can get other values for the mean and variance of the limit by including additive and multiplicative constants. The amazing thing about this limit is that it holds regardless of what the specific distribution of X_t is; it needs only to have a finite mean and variance. This much more general result is called the central limit theorem. We end the second part of this course, I think fittingly, with a proof of this result:[4]

[4] There are many variations of the central limit theorem with weaker assumptions than I am giving here. For example, I assume that the moment generating function is finite everywhere, although this need not be true. More general results require mathematical machinery outside the scope of this course.

Theorem 1 (Central Limit Theorem) *Let $Y_1, \ldots, Y_n$ be a sequence of independent, identically distributed random variables such that the moment generating function $m_X(t)$ exists and is finite for all values of t. Then:*

$$\frac{\sum_i Y_i - n \cdot \mu}{\sqrt{n\sigma^2}} \to Y \sim N(0, 1).$$

Where μ is the mean of each Y_i and σ^2 is the variance of each Y_i.

Proof. Without loss of generality, let X_n be a similarly defined sequence of random variables but with the condition that X has zero mean and unit variance. The general result follows by defining:

$$Y_n = \sigma X_n + \mu.$$

Now, define:

$$Z_n = \sum_i X_i / \sqrt{n}.$$

We proceed by showing that the moment generating function of Z_n limits to the moment generating function of a unit normal. More specifically, we show that the log of the moment generating function of Z_n limits to the log of the normal moment generating function.[5] Also, define $S_n = \sum_i X_i$. Now, using the rules we have for moment generating functions we get:[6]

[5] The log of the moment generating function is called the cumulant generating function, but you don't need to remember that for now.

[6] I'll use $M_X(t)$ to be the moment generating function for any X_n since they are all the same.

$$M_{S_n}(t) = [M_X(t)]^n$$

And:

$$\begin{aligned}
M_{Z_n}(t) &= \left[M_X(t/\sqrt{n})\right]^n \\
\log\left(M_{Z_n}(t)\right) &= \log\left[M_X(t/\sqrt{n})\right]^n \\
&= n \cdot \log\left(M_X(t/\sqrt{n})\right)
\end{aligned}$$

We need to take the limit of this function as n goes to infinity. It will be easier to consider a limit as some quantity becomes very small; let $\Delta = 1/\sqrt{n}$ and we can then consider Δ limiting to zero.

$$\lim_{\Delta \to 0}\left[\frac{\log\left(M_X(\Delta \cdot t)\right)}{\Delta^2}\right] = ?$$

The limit of both the top and bottom go to 0, so we can use l'Hôpital's rule: take the derivative of the top and the bottom and then compute the limit again.

$$\lim_{\Delta \to 0}\left[\frac{\log\left(M_X(\Delta \cdot t)\right)}{\Delta^2}\right] = \lim_{\Delta \to 0}\left[\frac{\frac{M_X'(t\Delta)t}{M_X(t\Delta)}}{2\Delta}\right]$$

But again, the limit of both numerator and denominator are zero. Taking out a constant and using l'Hôpital's rule once again we get:

$$\lim_{\Delta \to 0}\left[\frac{\frac{M_X'(t\Delta)t}{M_X(t\Delta)}}{2\Delta}\right] = \frac{t^2}{\Delta} \times \lim_{\Delta \to 0}\left[\frac{M_X''(t\Delta)}{M_X(t\Delta) + t\Delta M_X'(t\Delta)}\right]$$

And now we can take the limit by computing the limit of the numerator and denominator of the limit:

$$\frac{t^2}{2} \times \lim_{\Delta \to 0}\left[\frac{M_X''(t\Delta)}{M_X(t\Delta) + t\Delta M_X'(t\Delta)}\right] = \frac{t^2}{2} \times \frac{M_X''(0)}{M_X(0) + tM_X'(0)}$$

The value $M(0)$ is always equal to 1. Also, $M_X'(0)$ is the mean of X which is zero and $M_X''(0)$ is equal to 1. So:

$$\begin{aligned}
\frac{t^2}{2} \times \frac{M_X''(0)}{M_X(0) + tM_X'(0)} &= \frac{t^2}{2} \times \frac{1}{1+0} \\
&= \frac{t^2}{2}
\end{aligned}$$

Putting this together we get:

$$\lim_{n \to \infty}\left[\log\left(M_{Z_n}(t)\right)\right] = \frac{t^2}{2}$$

This is the moment generating function of a standard normal distribution, and therefore completes the proof ■.