Beginning of March 15, 2021

0.1 Geometric Random Variables

Random experiment with repeated, independent Bernoulli trials with the probability of success p.

The **Geometric random variable** X describes the number of trials until (and including) the first success, and we write $X \sim G(p)$.

$R(X) = \{1, 2, 3, \dots\}$ with $P_X(x) = P(X = x) = (1-p)^{x-1}p$. The cumulative distribution function (cdf) is a piecewise, monotone increasing function with limit 1 but never reaches it.

Mean, Variance, and MGF

The mean of $G(p) = 1/p$. Intuitively, if each independent Bernoulli trial has as probability of success p, it takes $1/p$ trials to get one success.

$$\begin{aligned}
\mu_X = E[X] &= \sum_{x=1}^{\infty} xP(x) = \sum_{x=1}^{\infty} x(1p-)^{x-1}p = p\sum_{x=1}^{\infty} x(1-p)^{x-1} \\
&= p\sum_{x=1}^{\infty} -\frac{\mathrm{d}}{\mathrm{d}p}(1-p)^x = -p\frac{\mathrm{d}}{\mathrm{d}p}\left[\sum_{x=1}^{\infty}(1-p)^x\right] = -p\frac{\mathrm{d}}{\mathrm{d}p}\left[\frac{1}{1-(1-p)} - 1\right] \\
&= -p\frac{\mathrm{d}}{\mathrm{d}p}\frac{1}{p} = -p(-1/p^2) = \frac{1}{p}.
\end{aligned}$$

The variance of $G(p)$ is $(1-p)/p^2$.

$$\begin{aligned}
\sigma_X^2 &= \mathrm{var}[X] = E[X^2] - E[X]^2 \\
&= \sum_{x=1}^{\infty} x^2 P(x) - \frac{1}{p^2} \\
&= \sum_{x=1}^{\infty} x^2(1-p)^{x-1}p - \frac{1}{p^2} \\
&= p\sum_{x=1}^{\infty}\left[x(x-1)+x\right](1-p)^{x-2}(1-p) - \frac{1}{p^2} \\
&= p(1-p)\sum_{x=1}^{\infty} x(x-1)(1-p)^{x-2} + \sum_{x=1}^{\infty} x(1-p)^{x-1}p - \frac{1}{p^2} \\
&= p(1-p)\sum_{x=1}^{\infty}\frac{\mathrm{d}^2}{\mathrm{d}p^2}\left[(1-p)^x\right] + \frac{1}{p} - \frac{1}{p^2} \\
&= p(1-p)\frac{\mathrm{d}^2}{\mathrm{d}p^2}\left[\sum_{x=1}^{\infty}(1-p)^x\right] + \frac{1}{p} - \frac{1}{p^2} \\
&= \frac{1-p}{p^2}
\end{aligned}$$

The MGF is

$$\begin{aligned}
\varphi_X(t) = E[e^{tx}] &= \sum_{x=1}^{\infty} e^{tx} P(x) \\
&= \sum_{x=1}^{\infty} e^{tx}(1-p)^{x-1}p \\
&= p\sum_{x=1}^{\infty} e^{t(x-1)}e^t(1-p)^{x-1} \\
&= pe^t \sum_{x=1}^{\infty} \left[e^t(1-p)\right]^{x-1} \\
&= pe^t \frac{1}{1-e^t(1-p)} && \text{notice that } |e^t(1-p)| < 1 \text{ locally} \\
&= \frac{pe^t}{1-e^t(1-p)}.
\end{aligned}$$

0.2 Negative Binomial Random Variables

A generalization of the geometric random variable, not that of a binomial random variable. This describes the random experiment of a repeated, independent Bernoulli trials with probability of success p.
X describes the number of trials to get r successes. We write $X \sim NB(r,p)$.
It follows that $R(X) = \{r, r+1, \dots\}$, and

$$P_X(x) = P(X=x) = \binom{x-1}{r-1}p^{r-1}\,(1-p)^{(x-1)-(r-1)}p = \binom{x-1}{r-1}\,p^r(1-p)^{x-r}$$

since if we need precisely x trials to get r successes, then there must be precisely $r-1$ successes in the first $x-1$ trials, and the x^{th} trial itself must be a success.
Note that $X \sim NB(1,p) \iff X \sim G(p)$.

Mean, Variance, and MGF

The mean of $NB(r,p)$ is r/p. Let X_i be the number of trials to get another success after already having $i-1$ successes. Then $X_i \sim G(p)$, and more importantly the X_i's are independent. Therefore,

$$E[X] = E\left[\sum_{i=1}^{r} X_i\right] = \sum_{i=1}^{r} E[X_i] = \frac{r}{p}.$$

The variance of X can be computed similarly, and $\text{var}[X] = r(1-p)/p^2$.
The MGF is

$$\begin{aligned}
\varphi_X(t) = E[e^{tX}] &= E[\exp(t\textstyle\sum)] \\
&= \prod_{i=1}^{r} E[e^{tX_i}] \\
&= \left(\frac{pe^t}{1-e^t(1-p)}\right)^r.
\end{aligned}$$

0.3 Poisson Distribution

The Poisson distribution is similar to binomial distribution, but the probability of success becomes a rate applied to a continuum as opposed to discrete selections.

Let a time interval of length 1 be given. Previously, without the concept of continuum, we divide this interval into n subintervals of length $1/n$ and perform independent Bernoulli trails on each one. Let the assumption be that the mean total number of successes among these n Bernoulli trials is λ. Then it immediately follows that the probability of success for each subinterval is λ/n (so that $n \cdot \lambda/n = \lambda$).

We fix this λ but let $n \to \infty$, namely dividing the interval into finer and finer subintervals of length $1/n$. What does the limit mean? It means we are on a continuum $[0,1]$ where each $x \in [0,1]$ resembles a "Bernoulli trial", such that the mean/expected number of total successes for all $x \in [0,1]$ is λ. Of course, as $n \to \infty$, $\lambda/n \to \infty$ and the previous definitions no longer make sense. Therefore, the **Poisson** distribution can be interpreted as a limit that consists of *infinitely many Bernoulli trials.*

Done with the heuristics, now we begin from finite cases (Bernoulli trials) and start the approximation. Suppose we have n independent Bernoulli trails. This naturally gives rise to a binomial distribution with parameters n and λ/n. Let X_n be $B(n, \lambda/n)$. Then

$$\begin{aligned} P_{X_n}(x) &= \binom{n}{x}(\lambda/n)^x(1-\lambda/n)^{n-x} \\ &= \frac{n!}{(n-x)!x!}\frac{\lambda^x}{n^x}(1-\lambda/x)^{n-x} \\ &= \frac{\lambda^x}{x!}(1-\lambda/n)^{-x}\frac{n!}{(n-x)!n^x}(1-\lambda/n)^n. \end{aligned}$$

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

$$\begin{aligned} \lim_{n\to\infty}\ln(1-\lambda/n)^n &= \lim_{n\to\infty} n\ln(1=\lambda/n) \\ &= \lim_{n\to\infty}\frac{\ln(1-\lambda/n)}{1/n} \\ &\overset{H}{=} \lim_{n\to\infty}\frac{\lambda/n^2 \cdot 1/(1-\lambda/n)}{-1/n^2} = -\lambda. \end{aligned}$$

On the other hand,

$$\frac{n!}{(n-x)!n^x} = \frac{n(n-1)\dots(n-x+1)}{n^x} = \frac{n}{n}\cdot\frac{n-1}{n}\cdot\dots\cdot\frac{n-x+1}{n} \to 1 \text{ as } n \to \infty.$$

Therefore,

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

Notation wise, $X \sim Pr(\lambda)$. $R(X) = \{0,1,2,\dots\}$ and $P_X(x) = e^{-\lambda}\lambda^x/x!$. Notice that the probabilities add up to 1:

$$\sum_{x=0}^{\infty} e^{-\lambda}\frac{\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x}{x!} = e^{-\lambda}e^{\lambda} = 1.$$

Mean, Variance, and MGF

The mean is given by

$$\begin{aligned} \mu_X = E[X] &= \sum_{x=0}^{\infty} xP(x) \\ &= \sum_{x=0}^{\infty} xe^{-\lambda}\frac{\lambda^x}{x!} = e^{-\lambda}\sum_{x=1}^{\infty}\frac{x\lambda^x}{x!} \\ &= e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^x}{(x-1)!} = \lambda e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda}{x!} \\ &= \lambda e^{-\lambda}e^{\lambda} = \lambda. \end{aligned}$$

> **Remark**. Of course this makes sense! On a continuum with λ being the mean occurrence rate of an event, what else do you expect $E[X]$ to be but λ itself?

The variance is given by

$$\mathrm{Var}[X] = E[X^2] - \lambda^2 = \sum_{x=0}^{\infty} \frac{x^2 \lambda^x e^{-\lambda}}{x!} - \lambda^2 = \lambda.$$

> **Remark**. *A nice problem assigned in HW7 gives the identity $E[X^n] = \lambda E[(X+1)^{n-1}]$ for Poisson distributions. If we apply it to $E[X^2]$, we immediately have*
>
> $$\mathrm{Var}[X] = \lambda E[(X+1)] - E[X]^2 = \lambda(\lambda+1) - \lambda^2 = \lambda.$$
>
> Alternatively, we can again use the idea of "limit of Bernoulli/binomial distributions":
>
> $$\mathrm{Var}[X] = \lim_{n\to\infty} \mathrm{Var}[B(n, \lambda/n)] = \lim_{n\to\infty} n \cdot (\lambda/n) \cdot (1 - \lambda/n) = \lambda.$$

The MGF is given by

$$\varphi_X(t) = E[e^{tX}] = e^{\lambda}(e^t - 1).$$