Tutorial 1

1 The Geometric Random Variable

Suppose that independent trials, each having probability p of being a success, are performed until a success occurs. If we let X be the number of trials required until the first success, then X is said to be a geometric random variable with parameter p. Its probability mass function is given by

$$p(n) = P\{X = n\} = (1-p)^{n-1}p, \quad n = 1, 2, \cdots .$$

Note that

$$\sum_{n=1}^{\infty} p(n) = p \sum_{n=1}^{\infty} (1-p)^{n-1} = \lim_{n\to\infty} p \frac{1-(1-p)^n}{1-(1-p)} = 1.$$

Expectation of a Geometric Random Variable:

$$\begin{aligned}
E[X] &= \sum_{n=1}^{\infty} np(1-p)^{n-1} \\
&= p \sum_{n=1}^{\infty} nq^{n-1} \\
&= p \sum_{n=1}^{\infty} \frac{d}{dq}(q^n) \\
&= p \frac{d}{dq} \left(\sum_{n=1}^{\infty} q^n \right) \\
&= p \frac{d}{dq} \left(\frac{q}{1-q} \right) \\
&= \frac{p}{(1-q)^2} \\
&= \frac{1}{p}
\end{aligned}$$

where $q = 1 - p$.

Intuitively, the expected number of independent trials we need to perform until we obtain our first success equals the reciprocal of the probability that any one trial results in a success.

2 The Poisson Random Variable

A random variable X, taking on one of the values $0, 1, 2, \cdots$, is said to be a Poisson random variable with parameter λ, if for some $\lambda > 0$,

$$p(i) = P\{X = i\} = e^{-\lambda} \frac{\lambda^i}{i!}, \quad i = 0, 1, \cdots .$$

Note that

$$\sum_{i=0}^{\infty} p(i) = e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^{-\lambda} e^{\lambda} = 1.$$

due to that the Taylor expansion of e^x at $x_o = 0$ is $e^x = \sum_{n=0}^{\infty} \frac{x}{n!}$.

An important property of the Poisson random variable is that it may be used to approximate a binomial random variable when the binomial parameter n is large and p is small. To see this, suppose that X is a binomial random variable with parameters (n, p), and let $\lambda = np$, i.e. $p = \frac{\lambda}{n}$. Then from the definition of the PMF of the binomial random variable

$$\begin{aligned}
P\{X = i\} &= \frac{n!}{(n-i)!i!} p^i (1-p)^{n-i} \\
&= \frac{n!}{(n-i)!i!} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i} \\
&= \frac{n(n-1)\cdots(n-i+1)}{i!} \frac{\lambda^i}{n^i} \frac{(1-\lambda/n)^n}{(1-\lambda/n)^i} \\
&= \frac{n(n-1)\cdots(n-i+1)}{n^i} \frac{\lambda^i}{i!} \frac{(1-\lambda/n)^n}{(1-\lambda/n)^i}
\end{aligned}$$

Now, for n large and p small

$$(1 - \frac{\lambda}{n})^n \approx e^{-\lambda}, \quad \frac{n(n-1)\cdots(n-i+1)}{n^i} \approx 1, \quad (1 - \frac{\lambda}{n})^i \approx 1.$$

Hence, for n large and p small

$$P\{X = i\} \approx e^{-\lambda} \frac{\lambda^i}{i!}.$$

Expectation of a Poisson random variable:

$$\begin{aligned}
E[X] &= \sum_{i=0}^{\infty} i e^{-\lambda} \frac{\lambda^i}{i!} \\
&= \sum_{i=1}^{\infty} i e^{-\lambda} \frac{\lambda^i}{i!} \\
&= \sum_{i=1}^{\infty} \frac{e^{-\lambda} \lambda^i}{(i-1)!} \\
&= e^{-\lambda} \lambda \sum_{i=1}^{\infty} \frac{\lambda^{i-1}}{(i-1)!} \\
&= e^{-\lambda} \lambda \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \\
&= e^{-\lambda} \lambda e^{\lambda} \\
&= \lambda
\end{aligned}$$

The moment generating function of a Poisson distribution with parameter λ.

$$\begin{aligned}\phi(t) &= E\left[e^{tX}\right] \\ &= \sum_{n=0}^{\infty} \frac{e^{tn}e^{-\lambda}\lambda^n}{n!} \\ &= e^{-\lambda}\sum_{n=0}^{\infty}\frac{(e^t\lambda)^n}{n!} \\ &= e^{-\lambda}e^{\lambda e^t} \\ &= \exp\{\lambda(e^t-1)\}\end{aligned}$$

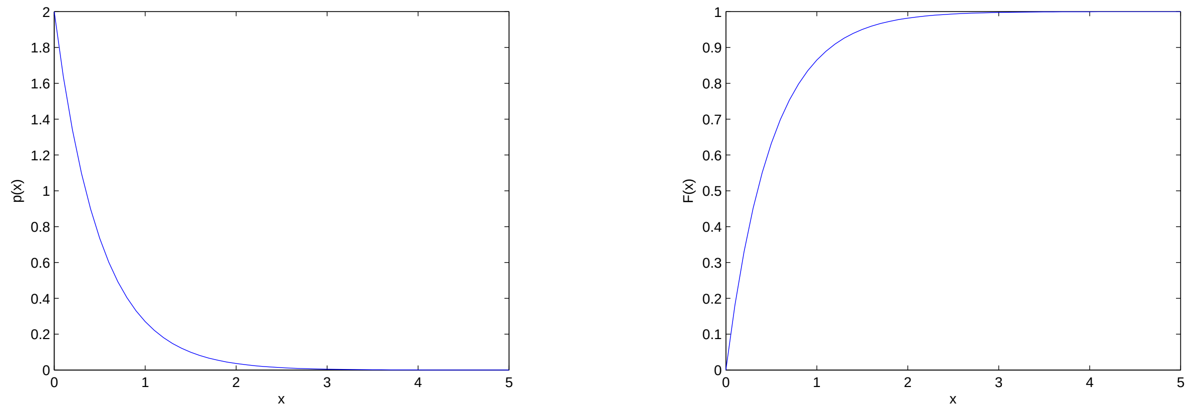
Differentiating yield

$$\begin{aligned}\phi'(t) &= \lambda e^t \exp\{\lambda(e^t-1)\}, \\ \phi''(t) &= (\lambda e^t)^2 \exp\{\lambda(e^t-1)\} + \lambda e^t \exp\{\lambda(e^t-1)\}\end{aligned}$$

and so

$$\begin{aligned}E[X] &= \phi'(0) = \lambda, \\ E[X^2] &= \phi''(0) = \lambda^2+\lambda, \\ \text{Var[X]} &= E[X^2]-(E[X])^2 = \lambda.\end{aligned}$$

3 Exponential Random Variables



(a) PDF

(b) CDF

Figure 1: An exponential random variable with parameter $\lambda = 2$

A continuous random variable whose probability density function is given, for some $\lambda > 0$, by $f(x) = \lambda e^{-\lambda x}$, if $x \geq 0$ and $f(x) = 0$ if $x < 0$ is said to be an exponential random variable with parameter λ. Note that the cumulative distribution function F is given by

$$F(a) = P(X \leq a) = \int_0^a \lambda e^{-\lambda x} dx = 1 - e^{-\lambda a}, \quad a \geq 0.$$

Note that $F(\infty) = \int_0^\infty \lambda e^{-\lambda x} dx = 1$.

The moment generating function of a exponential distribution with parameter λ.

$$\begin{aligned}
\phi(t) &= E[e^{tX}] \\
&= \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx \\
&= \lambda \int_0^\infty e^{-(\lambda - t)x} dx \\
&= \frac{\lambda}{\lambda - t} \quad t < \lambda.
\end{aligned}$$

Note by the preceding derivation that, for exponential distribution, $\phi(t)$ is only defined for value of t less than λ (otherwise $\phi(t) = \infty$). Differentiating of $\phi(t)$ yields

$$\begin{aligned}
\phi'(t) &= \frac{\lambda}{(\lambda - t)^2} \\
\phi''(t) &= \frac{2\lambda}{(\lambda - t)^3}
\end{aligned}$$

Hence

$$E[X] = \phi'(0) = \frac{1}{\lambda} \qquad E[X^2] = \phi''(0) = \frac{2}{\lambda^2}.$$

The variance is thus given by

$$\mathrm{Var(X)} = \mathrm{E[X^2]} - (\mathrm{E[X]})^2 = \frac{1}{\lambda^2}.$$

4 Properties of Covariance

For any random variables X, Y, Z and constant c,

1. $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$,
2. $\mathrm{Cov}(X, c) = 0$,
3. $\mathrm{Cov}(cX, Y) = c\mathrm{Cov}(Y, X)$,
4. $\mathrm{Cov}(X, Y + Z) = \mathrm{Cov}(X, Y) + \mathrm{Cov}(X, Z)$.

5 Take-home Exercise

Suppose we have random variables $X \sim \mathrm{Bern}(\frac{1}{2})$ and $Y \sim \mathrm{Bern}(\frac{1}{2})$, while X are Y are independent. Let us define a random variable Z as $Z = X \oplus Y$. Show that X, Y and Z are pairwise independent.

Proof:

We will first show that X and Z are independent, and a similar conclusion is straightforward due to the symmetry of this problem.

Note that

$$P(X=0)=P(X=1)=\frac{1}{2}, \qquad P(Y=0|Y=0)=\frac{1}{2}$$

According to the definition of Z, the possible values of Z are 0 and 1. Since X and Y are independent, we have

$$\begin{aligned}
P(Z=0,X=0)&=P(Y=0,X=0)=P(Y=0)P(X=0)=\frac{1}{4},\\
P(Z=0,X=1)&=P(Y=1,X=1)=P(Y=1)P(X=1)=\frac{1}{4},\\
P(Z=1,X=0)&=P(Y=1,X=0)=P(Y=1)P(X=0)=\frac{1}{4},\\
P(Z=1,X=1)&=P(Y=0,X=1)=P(Y=0)P(X=1)=\frac{1}{4}.
\end{aligned}$$

Thus, we have

$$\begin{aligned}
P(Z=0)&=\sum_x P(Z=0,X=x)=\frac{1}{2},\\
P(Z=1)&=\sum_x P(Z=1,X=x)=\frac{1}{2}.
\end{aligned}$$

So we have

$$\begin{aligned}
P(Z=0,X=0)&=P(Z=0)P(X=0)=\frac{1}{4},\\
P(Z=0,X=1)&=P(Z=0)P(X=1)=\frac{1}{4},\\
P(Z=1,X=0)&=P(Z=1)P(X=0)=\frac{1}{4},\\
P(Z=1,X=1)&=P(Z=1)P(X=1)=\frac{1}{4}.
\end{aligned}$$

Therefore, we have shown that X and Y are independent. □

6 Exercises

1. *Suppose that $X_1,\cdots,X_n$ are i.i.d random variables with expected value μ and variance σ^2. Let us define the sample mean as $\bar{X}=\sum_{i=1}^n X_i/n$. Show that*

$$\begin{aligned}
&(a)\quad E[\bar{X}]=\mu.\\
&(b)\quad \operatorname{Var}(\bar{X})=\frac{\sigma^2}{n}.\\
&(c)\quad \operatorname{Cov}(\bar{X},X_i-\bar{X})=0, i=1,\cdots,n.
\end{aligned}$$

Proof:

(a)

$$E[\bar{X}] = E\left[\sum_{i=1}^{n} \frac{X_i}{n}\right] = \frac{1}{n}\sum_{i=1}^{n} E\left[X_i\right] = \frac{n\mu}{n} = \mu$$

(b)

$$\mathrm{Var}(\bar{X}) = \mathrm{Var}\left(\sum_{i=1}^{n} \frac{X_i}{n}\right) = \left(\frac{1}{n}\right)^2 \mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} \mathrm{Var}\left(X_i\right) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

(c)

$$\begin{aligned}
\mathrm{Cov}(\bar{X}, X_i - \bar{X}) &= \mathrm{Cov}(\bar{X}, X_i) - \mathrm{Cov}(\bar{X}, \bar{X}) \\
&= \mathrm{Cov}\left(\frac{X_i + \sum_{j\neq i} X_j}{n}, X_i\right) - \mathrm{Var}(\bar{X}) \\
&= \frac{1}{n}\mathrm{Cov}\left(X_i + \sum_{j\neq i} X_j, X_i\right) - \frac{\sigma^2}{n} \\
&= \frac{1}{n}\mathrm{Cov}\left(X_i, X_i\right) + \frac{1}{n}\mathrm{Cov}\left(\sum_{j\neq i} X_j, X_i\right) - \frac{\sigma^2}{n} \\
&= \frac{1}{n}\mathrm{Var}(X_i) + \frac{1}{n}\sum_{j\neq i}\mathrm{Cov}\left(X_i, X_j\right) - \frac{\sigma^2}{n} \\
&= \frac{\sigma^2}{n} - 0 - \frac{\sigma^2}{n} \\
&= 0
\end{aligned}$$

□

2. *Suppose we know that the number of items produced in a factory during a week is a random variable with mean* 500.

(a) What is the probability that this week's production will be at least 1000*?*

(b) If the variance of a week's production is known to equal 100*, then what can be said about the probability that this week's production will between* 400 *and* 600*?*

Solution:

Let X be the number of items that will be produced in a week.

(a) By Markov's inequality,

$$P(X \geq 1000) \leq \frac{E[X]}{1000} = \frac{500}{1000} = \frac{1}{2}.$$

(b) By Chebyshev's inequality,

$$P(|X - 500| \geq 100) \leq \frac{\sigma^2}{100^2} = \frac{1}{100}.$$

Hence,

$$P(|X - 500| \leq 100) \geq 1 - \frac{1}{100} = \frac{99}{100}$$

and so the probability that this week's production will be between 400 and 600 is at least 0.99. $\square$

Notes:

a. (**Markov's Inequality**) If X is a random variable that takes only nonnegative values, then for any value $a > 0$, $P(X \geq a) \leq \frac{E[X]}{a}$.

b. (**Chebyshev's Inequality**) If X is a random variable with mean μ and variance σ^2, then, for any value $k > 0$, $P(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}$.

3. *Let* X_1, X_2, $\cdots$, X_{10} *be independent random variables, each being uniformly distributed over* $(0, 1)$. *Estimate* $P(\sum_1^{10} X_i > 7)$.

Solution:

Since $E[X_i] = \frac{1}{2}$, $\mathrm{Var(X_i)} = \frac{1}{12}$, we have by the central limit theorem that

$$\begin{aligned} P\left(\sum_1^{10} X_i > 7\right) &= 1 - P\left(\sum_1^{10} X_i \leq 7\right) \\ &= 1 - P\left(\frac{\sum_1^{10} X_i - 10\left(\frac{1}{2}\right)}{\sqrt{10\left(\frac{1}{12}\right)}} \leq \frac{7-5}{\sqrt{10\left(\frac{1}{12}\right)}}\right) \\ &\approx 1 - \Phi(2.2) \\ &= 0.0139 \end{aligned}$$

Notes:

a. For a uniform random variable over (a, b), the mean is $\frac{a+b}{2}$, and its variance is $\frac{(b-a)^2}{12}$.

b. (**Central Limit Theorem**) Let X_1, X_2, $\cdots$ be a sequence of i.i.d. random variables, each with mean μ and variance σ^2. Then the distribution of $\frac{X_1+X_2+\cdots+X_n-n\mu}{\sigma\sqrt{n}}$ tends to the standard normal as $n \to \infty$. That is $P\left(\frac{X_1+X_2+\cdots+X_n-n\mu}{\sigma\sqrt{n}} \leq a\right) \to \Phi(a)$ as $n \to \infty$.

4. *The lifetime of a special type of battery is a random variable with mean* 40 *hours and standard deviation* 20 *hours. A battery is used until it fails, at which point it is replaced by a new one. Assuming a stockpile of* 25 *such batteries, the lifetime of which are independent, approximate the probability that over* 1100 *hours of use can be obtained.*

Solution:

If we let X_i denote the lifetime of the ith battery to be put in use, then we desire $p = P(X_1 + \cdots + X_{25} > 1100)$, which is approximated as follows:

$$\begin{aligned} p &= P(X_1 + \cdots + X_{25} > 1100) \\ &= 1 - P(X_1 + \cdots + X_{25} \leq 1100) \\ &= 1 - P(\frac{X_1 + \cdots + X_{25} - 25 \times 40}{20\sqrt{25}} \leq \frac{1100 - 40 \times 25}{20\sqrt{25}}) \\ &\approx 1 - \Phi(1) \\ &\approx 0.1587 \end{aligned}$$

$\square$