Chap 2 Note: Random Variables

1 Random Variables and CDF

Example Suppose that independent trials, each of which results in any of m possible outcomes with respective probabilities $p_1, p_2, \cdots, p_m$, $\sum_{i=1}^{m} p_i = 1$, are continually performed. Let X denote the number of trials needed until each outcome has occurred at least once.

Explain Rather than directly considering $P\{X = n\}$ we will first determine $P\{X > n\}$, the probability that at least one of the outcomes has not yet occurred after n trails. Letting A_i denote the event that outcome i has not yet occurred after the first n trails, $i = 1, \cdots, m$, then

$$
\begin{aligned}
P\{x > n\} =& P(\cup_{i=1}^{m} A_i) \\
=& \sum_{i=1}^{m} P(A_i) - \sum\sum_{i<j} P(A_i A_j) + \cdots \\
&+ \sum\sum\sum_{i<j<k} P(A_i A_j A_k) - \cdots + (-1)^{m+1} P(A_1 \cdots A_n)
\end{aligned}
$$

$P(A_i)$ is the probability that each of the first n trails results in a non-i outcome, and so by independence $P(A_i) = (1 - p_i)^n$

Similarly, $P(A_i A_j)$ is the probability that the first n trails all result in a non-i and non-j outcome, and so

$$P(A_i A_j) = (1 - p_i - p_j)^n$$

As all of the other probabilities are similar, we see that

$$
\begin{aligned}
P\{X > n\} =& \sum_{i=1}^{m} (1 - p_i)^n - \sum\sum_{i<j} (1 - p_i - p_j)^n \\
&+ \sum\sum\sum_{i<j<k} (1 - p_i - p_j - p_k)^n - \cdots
\end{aligned}
$$

Since $P\{X = n\} = P\{X > n - 1\} - P\{x > n\}$, we see, upon using the algebraic identity $(1 - a)^{n-1} - (1 - a)^n = a(1 - a)^{n-1}$, that

$$
\begin{aligned}
P\{X = n\} =& \sum_{i=1}^{m} p_i * (1 - p_i)^{n-1} - \sum\sum_{i<j} (p_i + p_j) * (1 - p_i - p_j) \\
&+ \sum\sum\sum_{i<j<k} (p_i + p_j + p_k)(1 - p_i - p_j - p_k)^{n-1} - \cdots
\end{aligned}
$$

CDF The cumulative distribution function(cdf) $F(.)$ of the random variable X is defined for any real number b, $-\infty < b < \infty$, by

$$F(b) = P\{X \leq b\}$$

We have

$$P\{a < X \leq b\} = F(b) - F(a), \text{for all } a < b$$
$$P\{X < b\} = \lim_{h \to 0^+} P\{X \leq b - h\} = \lim_{h \to 0^+} F(b - h)$$

2 Discrete Random Variables

2.1 The Bernoulli Random Variable

Suppose that a trail, or an experiment, whose outcome can be classified as either a "success" or as a "failure" is performed. If we let X equal 1 if the outcome is a success and 0 if it is a failure, then the probability mass function of X is given by

$$p(0) = P\{X = 0\} = 1 - p$$
$$p(1) = P\{X = 1\} = p$$

where $p, 0 \leq p \leq 1$, is the probability that the trail is a "success".

A random variable X is said to be a *Bernoulli random variable* if its probability mass function is given by the above equation for $p \in (0, 1)$.

2.2 The Binomial Random Variable

Suppose that n independent trails, each of which results in a "success" with probability p and in a "failure" with probability $1 - p$, are to be performed. If X represents the number of successes that occur in the n trails, then X is said to be a *binomial random variable* with parameters (n, p).

The probability mass function of a binomial random variable having parameters (n, p) is given by

$$p(i) = \binom{n}{i} p^i * (1 - p)^{n-i}, i = 0, 1, \cdots, n$$

where

$$\binom{n}{i} = \frac{n!}{(n - i)! * i!}$$

Note that, by the binomial theorem, the probabilities sum to one, that is,

$$\sum_{i=0}^{\infty} p(i) = \sum_{i=0}^{n} p^i * (1 - p)^{n-i} = (1 + (1 - p))^n = 1$$

2.3 The Geometric Random Variable

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

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

To check that $p(n)$ is a probability mass function, we note that

$$\sum_{n=1}^{\infty} p(n) = \sum_{n=1}^{\infty} (1 - p)^{n-1} = 1$$

2.4 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 $\lambda \geq 0$,

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

The above equation defines a probability mass function since

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

Approximate a Binomial Random Variable by Poisson An important property of the Poisson random variable is that it may e 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 $\lambda = n * p$. Then

$$\begin{aligned}
PX = i =& \frac{n!}{(n-i)! * i!} p^i * (1-p)^{n-i} \\
=& \frac{n!}{(n-i)! * i!} * (\lambda/n)^i (1-\lambda/n)^{n-i} \\
=& \frac{n * (n-1) * \cdots * (n-i+1)}{n^i} * \frac{\lambda^i}{i!} * \frac{(1-\lambda/n)^n}{(1-\lambda/n)^i}
\end{aligned}$$

For n large and p small,

$$\begin{aligned}
&(1-\lambda/n)^n \approx e^{-\lambda} \\
&\frac{n * (n-1) * \cdots * (n-i+1)}{n^i} \approx 1 \\
&(1-\lambda/n)^i \approx 1
\end{aligned}$$

Hence for n large and p small, we have

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

3 Continuous Random Variables

Probability Density Function $f(x)$ is called the probability density function, which is a derivative of cumulative distribution function(CDF). We have

$$\begin{aligned}
&P\{a \leq X \leq b\} = \int_a^b f(x)dx \\
&P\{X = a\} = \int_a^a f(x)dx = 0 \\
&P\{a - \epsilon/2 \leq X \leq a + \epsilon/2\} = \int_{a-\epsilon/2}^{a+\epsilon/2} f(x)dx \approx \epsilon * f(a)
\end{aligned}$$

In other words, the probability that X will be contained in an interval of length ϵ around the point a is approximately $\epsilon * f(a)$. From this, we see that $f(a)$ is a measure of how likely it is that the random variable will be near a.

3.1 The Uniform Random Variable

A random variable is said to be uniformly distributed over the interval $(0, 1)$ if its probability density is given by

$$f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & otherwise. \end{cases}$$

Note that the preceding is a density function since $f(x) \geq 0$ and

$$\int_{-\infty}^{\infty} f(x)dx = \int_{0}^{1} = 1$$

In general, we say that X is a uniform random variable on the interval (α, β) if its probability density function is given by

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha} & \text{if } \alpha < x < \beta \\ 0 & otherwise \end{cases}$$

3.2 Exponential Random Variable

A continuous random variable whose probability density function is given, for some $\lambda > 0$, by

$$f(x) = \begin{cases} \lambda * e^{-\lambda * x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

is said to be an exponential random variable with parameter λ.

3.3 Gamma Random Variables

A continuous random variable whose density is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} (\lambda x)^{\alpha - 1} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

for some $\lambda > 0, \alpha > 0$ is said to be a gamma random variable with parameter α, λ. The quantity $\Gamma(\alpha)$ is called the gamma function and is defined by

$$\Gamma(\alpha) = \int_{0}^{\infty} e^{-x} x^{\alpha - 1} dx$$

It is easy to show by induction that for integral α, say $\alpha = n$,

$$\Gamma(n) = (n - 1)!$$

3.4 Normal Random Variables

We say that X is a normal random variable(or simply that X is normal distributed) with parameters μ and σ^2 of X is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}, -\infty < x < \infty$$

The density function is a bell-shaped curve that is symmetric around μ.

An important fact about normal random variables is that if X is normally distributed with parameter μ and σ^2 then $Y = \alpha * X + \beta$ is normally distributed with parameters $\alpha * \mu + \beta$ and $\alpha^2 \sigma^2$.

Proof. Suppose first that $\alpha > 0$ and note that $F_Y(.)$, the cumulative distribution function of the random variable Y, is given by

$$
\begin{aligned}
F_Y(a) =& P\{Y \leq a\} \\
=& P\{\alpha X + \beta \leq a\} \\
=& P\{X \leq \frac{a-\beta}{\alpha}\} \\
=& F_X(\frac{a-\beta}{\alpha}) \\
=& \int_{-\infty}^{(a-\beta)/\alpha} \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} dx \\
=& \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}\alpha\sigma} exp\{\frac{-(v-(\alpha\mu+\beta))^2}{2\alpha^2\sigma^2}\} dv
\end{aligned}
$$

where the last equality is obtained by the change in variables $v = \alpha x + \beta$. However, since $F_Y(a) = \int_{-\infty}^{a} f_Y(v)dv$, it follows from the Equation that the profit density function $f_Y(.)$ is given by

$$
f_Y(v) = \frac{1}{\sqrt{2\pi}\alpha\sigma} exp\{\frac{-(v-(\alpha\mu+\beta))^2}{2(\alpha\sigma)^2}\}, -\infty < v < \infty
$$

Hence, Y is normally distributed with parameters $\alpha\mu + \beta$ and $(\alpha\mu)^2$. A similar result is also true when $\alpha < 0$. □

Implementation One implementation of the preceding result is that if X is normally distributed with parameters μ and σ^2 then $Y = \frac{X-\mu}{\sigma}$ is normally distributed with parameters 0 and 1. Such a random variable Y is said to have the standard or unit normal distribution.

4 Expectation of a Random Variable