Probability

Part 3 - Random Variables, Distributions and Expected Values

1. References

(1) R. Durrett, *The Essentials of Probability*, Duxbury.

(2) L.L. Helms, *Probability Theory with Contemporary Applications*, Freeman.

(3) J.J. Higgins and S. Keller-McNulty, *Concepts in Probability and Stochastic Modeling*, Duxbury.

(4) R.V. Hogg and E.A. Tanis, *Probability and Statistical Inference*, Prentice-Hall.

2. Solved Problems

Problem 1. A fair die is tossed. The random variable X is equal to 1 if a one or a six occurs and is equal to 0 otherwise. Find the probability mass function of X

Solution:

A random variable X is a function $X : S \longrightarrow \mathbb{R}$ from the sample space of a random experiment to the real numbers. In this case, $S = \{1, \cdots, 6\}$ and $X(1) = X(6) = 1$ while $X(s) = 0$ for all other $s \in S$. The only restriction on such a function is that if I is an interval of real numbers, the set of outcomes

$$[X \in I] = \{s \in S | X(s) \in I]$$

must be an event. For finite or countable sample spaces S this is no restriction and any real valued function on S is a random variable. For uncountable sample spaces there may be functions that are not random variables, but they are hard to find. Any "ordinary" function is a random variable.

A discrete random variable is one that has only finitely many or countably many distinct values. For a discrete random variable X, the probability mass function is defined for real numbers x as $f(x) = P[X = x]$. In the present case, $f(1) = 2/6$, $f(0) = 4/6$, and $f(x) = 0$ for any other value of x.

A random variable whose only values are 0 and 1 is called a Bernoulli random variable. Bernoulli random variables are often used to encode the occurrence or nonoccurrence of an event. The event $[X = 1]$ is often designated a "success" and its complement $[X = 0]$ is a "failure". The probability $p = P[X = 1]$ is called the success probability. For a Bernoulli random variable X with success probability p, the probability mass function is $f(x) = p^x(1-p)^{1-x}$ for $x = 0, 1$ and $f(x) = 0$ for all other x.

Problem 2: Let X be a Bernoulli random variable with success probability p. Find the mean, variance, and standard deviation of X.

Solution:

The expected value or expectation of a discrete random variable X is defined as

$$E[X] = \sum_x xP[X = x] = \sum_x xf(x)$$

Symbolically, the sum may be taken over all real numbers x, but since $f(x) = 0$ for all but countably many values of x, it actually reduces to a finite sum or an infinite series. If it is an infinite series, it must be absolutely convergent, otherwise the expected value does not exist. In the present case, we sum over the only two possible values $x = 0$ and $x = 1$. Thus,

$$\begin{aligned} E[X] &= 1 \cdot P[X = 1] + 0 \cdot P[X = 0] \\ &= 1 \cdot p + 0 \cdot (1 - p) = p \end{aligned}$$

The expected value of a random variable is often denoted with the Greek letter μ.

The variance of a random variable X is defined to be

$$var(X) = E[(X - E[X])^2] = E[X^2] - E[X]^2$$

if the expected value exists. In the present case, $X^2 = X$, so $E[X^2] = E[X] = p$ and $E[X]^2 = p^2$. Therefore, $var(X) = p - p^2 = p(1 - p)$.

The square root of the variance is called the standard deviation and is often denoted with the Greek letter σ. Therefore, $var(X) = \sigma^2$. Here, $\sigma = \sqrt{p(1 - p)}$.

We remark that if X is a discrete random variable and $g : \mathbb{R} \longrightarrow \mathbb{R}$ is a reasonable function, the expected value of the transformed random variable $Y = g(X)$ may be found by

$$E[Y] = E[g(X)] = \sum_x g(x)f(x)$$

provided the sum is absolutely convergent. This is rather snidely called the law of the unconscious statistician.

Problem 3: Suppose a coin is biased and has head probability p. Toss the coin n times and let $X_i = 1$ if the i^{th} toss is a head and $X_i = 0$ if it is a tail. Assume that the random variables $X_1, \cdots, X_n$ are independent. Find their joint probability mass function.

Solution:

The random variables $X_1, \cdots, X_n$ are *jointly distributed* because they arise from the same random experiment,i.e., they are defined on the same sample space. They are independent if for each sequence $I_1, \cdots, I_n$ of intervals, the events $[X_i \in I_i]$ are independent events, i.e.,

$$P[X_1 \in I_1; \cdots ; X_n \in I_n] = P[X_1 \in I_1] \cdots P[X_n \in I_n].$$

Since these are discrete random variables, this is equivalent to

$$P[X_1 = x_1; X_2 = x_2; \cdots ; X_n = x_n] = \prod_{i=1}^{n} P[X_i = x_i]$$

for each sequence of possible values (0's and 1's, in this case). The left hand side of this equation is called the joint probability mass function of the X_i and is commonly denoted by $f(x_1, \cdots, x_n)$. Denoting the individual (or marginal) probability mass function of X_i by f_i, $X_1, \cdots, X_n$ are independent if and only if

$$f(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} f_i(x_i)$$

for all real numbers $x_1, \cdots, x_n$. In this case, if $x_i = 0$ or $x_i = 1$, $f_i(x_i) = p^{x_i}(1-p)^{1-x_i}$. Thus,

$$f(x_1, x_2, \cdots, x_n) = p^y(1-p)^{n-y}$$

where $y = \sum x_i$.

Problem 4: Two cards are drawn without replacement from a standard deck. All outcomes are equally likely. Let $X_i = 1$ if the i^{th} card drawn is a heart and $X_i = 0$ if it is not. Find the joint pmf of X_1 and X_2. Find their marginal pmf's. Show that X_1 and X_2 are dependent.

Solution:

X_1 and X_2 are Bernoulli variables. The success probability for X_1 is $P[X_1 = 1] = P[\text{Heart on first draw}] = 1/4$. Therefore, $f_1(x_1) = (\frac{1}{4})^{x_1}(\frac{3}{4})^{1-x_1}$ if $x_1 = 1$ or $x_1 = 0$. It is left to you to show that the success probability for X_2 is

$$P[X_2 = 1] = P[X_2 = 1|X_1 = 0]P[X_1 = 0] + P[X_2 = 1|X_1 = 1]P[X_1 = 1] = 1/4$$

also. Therefore, X_1 and X_2 have the same Bernoulli distribution. The product of the marginal pmf's is $f_1(1)f_2(1) = (\frac{1}{4})^2$, $f_1(1)f_2(0) = f_1(0)f_2(1) = (\frac{1}{4})(\frac{3}{4})$, $f_1(0)f_2(0) = (\frac{3}{4})^2$. On the other hand, $f(1,1) = \frac{13 \cdot 12}{52 \cdot 51}$. This is enough to show that X_1 and X_2 are dependent. Continuing, $f(1,0) = f(0,1) = (\frac{1}{4})(\frac{39}{51})$ and $f(0,0) = (\frac{3}{4})(\frac{38}{51})$.

Problem 5: In the preceding problem, let $Y = X_1 + X_2$ be the total number of hearts drawn. Find the expected value and variance of Y.

Solution: The expected value operator "E" is linear. Thus, $E[Y] = E[X_1 + X_2] = E[X_1] + E[X_2] = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$.

The variance is not linear except in special circucumstances. Here it can be calculated directly as follows.

$$\begin{aligned} var(Y) &= E[Y^2] - E[Y]^2 \\ &= E[X_1^2 + 2X_1X_2 + X_2^2] - \frac{1}{4} \\ &= E[X_1^2] + 2E[X_1X_2] + E[X_2^2] - \frac{1}{4} \\ &= \frac{1}{4} + \frac{1}{4} - \frac{1}{4} + 2E[X_1X_2] \\ &= \frac{1}{4} + 2E[X_1X_2] \end{aligned}$$

To calculate the second term of the last expression, it helps to remember that X_1X_2 is itself a Bernoulli variable, with success probability $f(1,1) = \frac{3}{51}$. Therefore, $E[X_1X_2] = \frac{1}{17}$ and $var(Y) = \frac{1}{4} + \frac{2}{17} = \frac{25}{68}$.

When the terms of a sum of random variables are independent, the variance is linear. Thus, if the cards in this example were drawn with replacement, the variance of Y would be $var(X_1) + var(X_2) = \frac{3}{8}$.

Problem 6: Let $X_1, X_2, \cdots, X_n$ be independent Bernoulli random variables with common success probability p. Let $Y = X_1 + X_2 + \cdots + X_n$. Find the pmf of Y, the expected value of Y, and the variance of Y.

Solution: The possible values of Y are the integers from 0 to n inclusive. Let y denote an arbitrary one of these values. By Problem 3, the probability of any particular sequence of y 1's and $n - y$ 0's is $p^y(1-p)^{n-y}$. There are $\binom{n}{y}$ such sequences of 0's and 1's. Therefore, $P[Y = y] = \binom{n}{y} p^y(1-p)^{n-y}$.

By the linearity of expectation,

$$E[Y] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p = np,$$

and because the X_i are independent (see the preceding problem),

$$var(Y) = \sum_{i=1}^{n} var(X_i) = np(1-p).$$

This distribution is called the *binomial distribution* with n trials and success probability p.

Problem 7: A random variable T has an exponential distribution. Let t_1 and t_2 be arbitrary positive numbers. Show that

$$P[T > t_1 + t_2 | T > t_1] = P[T > t_2].$$

Solution: Unlike the distributions considered up to now, an exponential distribution is not discrete. Rather, it is continuous and instead of a probability mass function it has a probability density function $f(t)$. This is a nonnegative function such that

$$P[T \in I] = \int_I f(t)dt.$$

for each interval I. In particular, $P[T > \tau] = \int_\tau^\infty f(t)dt$. For an exponential distribution the density function is

$$f(t) = \frac{1}{\mu} e^{-\frac{t}{\mu}}$$

for $t \geq 0$ and $f(t) = 0$ for $t < 0$. The parameter μ is an arbitrary but fixed positive number. For $\tau > 0$,

$$
\begin{aligned}
P[T > \tau] &= \int_{\tau}^{\infty} \frac{1}{\mu} e^{-\frac{t}{\mu}} dt \\
&= e^{-\frac{\tau}{\mu}}.
\end{aligned}
$$

Now,

$$
\begin{aligned}
P[T > t_1 + t_2 | T > t_1] &= P[T > t_1 + t_2]/P[T > t_1] \\
&= e^{-(t_1+t_2)}/e^{-t_1} \\
&= e^{-t_2} = P[T > t_2]
\end{aligned}
$$

If T is the lifetime of a system (say an ionized atom of Hydrogen), the conclusion above says that given that the system has survived past time t_1, the probability that it survives an additional time t_2 is the same as the unconditional probability that it survives for time t_2 from the beginning. Such system lifetimes are called *memoryless* and it can be shown that the only continuous distributions that have this property are the exponential distributions with various values of μ.

Problem 8: Show that the expected value of a random variable with an exponential distribution is the parameter μ.

Solution:

For a continuous random variable, such as T, with a density function, the expected value is defined as

$$E[T] = \int_{-\infty}^{\infty} f(t)dt.$$

More generally, if g is a reasonable function from $\mathbb{R}$ to $\mathbb{R}$,

$$E[g(T)] = \int_{-\infty}^{\infty} g(t)f(t)dt$$

provided these integrals are absolutely convergent. A "reasonable" function is, to be precise, a measureable function. Any function that can be obtained as the pointwise limit of a sequence of continuous functions is measureable.

Applying this to an exponential random variable,

$$E[T] = \int_{0}^{\infty} t\frac{1}{\mu}e^{-\frac{t}{\mu}}dt.$$

Integration by parts gives $E[T] = \mu$. To find the variance, we use the formula $var(T) = E[T^2] - E[T]^2$.

$$E[T^2] = \int_{0}^{\infty} t^2\frac{1}{\mu}e^{-\frac{t}{\mu}}dt.$$

Again by parts this is $E[T^2] = 2\mu^2$. Thus, $var(T) = 2\mu^2 - \mu^2 = \mu^2$, and the standard deviation is μ.

Problem 9: Let Y have a binomial distribution with $n = 100$ trials and success probability $p = .2$. Find the approximate value of the probability $P[16 < Y \leq 28]$.

Solution:

The random variable Y has the same distribution as $\sum_{i=1}^{100} X_i$, where $X_1, \cdots, X_{100}$ are independent Bernoulli random variables with success probability .2. Therefore, let us write $Y = \sum_{i=1}^{100} X_i$. We will *standardize* Y by subtracting its mean value and dividing by its standard deviation and calling the resulting random variable Z.

$$Z = \frac{Y - E[Y]}{\sqrt{var(Y)}} = \frac{Y - np}{\sqrt{np(1-p)}} = \frac{Y - 20}{4}$$

The Central Limit Theorem asserts that for sums, such as Y, of independent and identically distributed random variables, the distribution of the standardized sum approaches the standard normal distribution as the number n of summands grows without bound. The standard normal distribution is a continuous distribution with density function

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}.$$

The Central Limit Theorem is the reason why the normal distributions, in particular the standard normal distribution, are so important in probability theory.

Applying the Central Limit Theorem,

$$\begin{aligned} P[16 < Y \leq 28] &= P[-1 < \frac{Y - 20}{4} \leq 2] \\ &= P[-1 < Z \leq 2] \approx \int_{-1}^{2} \phi(z) dz \end{aligned}$$

This integral cannot be evaluated by elementary calculus, but there are extensive tables of integrals of the standard normal density function. According to these tables, the numerical answer is 0.8186. The true answer is 0.7876. The accuracy of the Central Limit approximation could be increased by modifying the interval to avoid integer endpoints. For example, the event $[16 < Y \le 28]$ is the same as the event $[16.5 < Y \le 28.5]$. With this correction, the Central Limit approximation is 0.7924.

Problem 10: Let Z be a random variable with the standard normal distribution. Let $\mu \in \mathbb{R}$ and $\sigma > 0$ be arbitrary. Define a new random variable X by $X = \sigma Z + \mu$. Find the density function of X.

Solution:

Let $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$ be the standard normal density function and let $\Phi(\zeta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\zeta} e^{-\frac{1}{2}z^2} dz$ denote its cumulative distribution function. We will find the cumulative distribution function of X and, by differentiation, the density function of X. The cumulative distribution function is

$$
\begin{aligned}
F(x) = P[X \le x] = \int_{-\infty}^{x} f(u) du \\
= P[\sigma Z + \mu \le x] \\
= P[Z \le \frac{x-\mu}{\sigma}] \\
= \Phi(\frac{x-\mu}{\sigma}).
\end{aligned}
$$

Differentiate both sides with respect to x. On one side we get $f(x)$. On the other side, we get

$$\frac{1}{\sigma}\phi(\frac{x-\mu}{\sigma}) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}.$$

This is the normal density function with mean μ and standard deviation σ.

3. Unsolved Problems

Problem 1: Let X have a binomial distribution with n trials and success probability p. Let t be an arbitrary real number. Find $E[e^{tX}]$. Hint: Use the binomial theorem.

Problem 2: If X_1 and X_2 are jointly distributed random variables, their *covariance* is defined as $cov(X_1, X_2) = E[X_1X_2] - E[X_1]E[X_2]$. Draw two cards in succession from a standard deck Let $X_i = 1$ if the i^{th} draw is a heart and let $X_i = 0$ otherwise. Find the covariance between X_1 and X_2 if the cards are drawn with replacement.

Problem 3: In the preceding problem, find the covariance when the cards are drawn without replacement, that is, independently.

Problem 4: A *Poisson* random variable X is a discrete random variable whose values are the nonnegative integers and whose probability mass function is

$$f(x) = e^{-\mu}\frac{\mu^x}{x!}$$

for $x = 0, 1, 2, \cdots$. Show that the expected value of X is μ.

Problem 5: Forty percent of voters in a large city would approve a new bond issue. A random sample of 100 voters is taken with replacement.

Approximately what is the probability that more than 50 of the sampled voters would approve the issue? You will need to find a table of the standard normal distribution. Every statistics book has one.

Problem 6: Let Z have the standard normal distribution. Show that $E[Z] = 0$ and $var(Z) = 1$.