Averages of Random Variables

Lecture 5

Spring 2002
Averages of Random Variables

Suppose that a random variable $U$ can take on any one of $L$ random values, say $u_1, u_2, \ldots, u_L$. Imagine that we make $n$ independent observations of $U$ and that the value $u_k$ is observed $n_k$ times, $k = 1, 2, \ldots, L$. Of course, $n_1 + n_2 + \cdots + n_L = n$. The empirical average can be computed by

$$\bar{U} = \frac{1}{n} \sum_{k=1}^{L} n_k u_k = \sum_{k=1}^{L} \frac{n_k}{n} u_k$$

The concept of statistical averages extends from this simple concept
Expected Value

The expected value of a discrete random variable $\mathcal{U}$ is defined by

$$E[\mathcal{U}] = \sum_{\mathcal{U}} u_k P(u_k)$$

Note how this definition compares with the empirical average.

The empirical average $\bar{\mathcal{U}}$ is a random variable. It will have a different value with each set of sample values.

The expected value is a single number. It is a parameter of the distribution. How do you think the expected value and the random variable $\bar{\mathcal{U}}$ relate?
Die Tossing Example

Emperical Average of Die Tosses

Average over 12 Trials -- 1000 Repetions

Emperical Average of Die Tosses

Average over 120 Trials -- 1000 Repetions
Function of a Discrete Random Variable

Let $U$ be a discrete random variable and let $V = g(U)$ Suppose that we want to compute the average value of $V$ rather than of $U$.

Given that $U$ can take on the values $u_k$ with probabilities $P[U = u_k]$, $k = 1, 2, \ldots, L$, we can compute both the values and probabilities that can be taken on by $V$.

Let $v_1, v_2, \ldots, v_r$ be the set of values that can be assumed by $V$ and let $P[V = v_j]$, $j = 1, 2, \ldots r$ be the corresponding probabilities. Then

$$E[V] = \sum_{j=1}^{r} v_j P[V = v_j]$$

We need to determine the values and the probabilities.
Expected Value of $V = g(U)$

By making use of the relationship $v = g(u)$ we can now rewrite this in a form that associates $g(u_k)$ with $P[U = u_k]$. The result is the nice compact equation

$$E[V] = \sum_{k=1}^{L} g(u_k) P[U = u_k]$$

If you were given a machine to generate samples of $U$, how would you go about estimating $E[V]$.
Continuous Random Variables

The expected value of a continuous random variable $U$ is

$$E[U] = \int_{-\infty}^{\infty} u f_U(u) du$$

The expected value of a function of a random variable $V = g(U)$ is

$$E[V] = \int_{-\infty}^{\infty} g(u) f_U(u) du$$

It is not necessary to compute $f_V(v)$
Example

Let $U$ be a random variable with an exponential distribution with parameter $a$.

$$f_U(u) = \begin{cases} 
  ae^{-au} & \text{for } u \geq 0 \\
  0 & \text{for } u < 0 
\end{cases}$$

To find the expected value of $U$ we calculate the integral

$$E[U] = \int_0^\infty uae^{-au}du = \frac{1}{a}$$

The parameter $a$ is therefore the reciprocal of the expected value. You should sketch the exponential distribution for a few values of $a$ and convince yourself that this is a reasonable result.
Example: $V = U^2$

Let $U$ have an exponential distribution with parameter $a$, and let $V = U^2$. Find $E[V] = E[U^2]$.

$$f_U(u) = \begin{cases} ae^{-au} & \text{for } u \geq 0 \\ 0 & \text{for } u < 0 \end{cases}$$

$$E[U^2] = \int_0^\infty u^2 ae^{-au} du = \frac{2}{a^2}$$
Mean, Variance and Standard Deviation

The mean value of a random variable $U$ is $\mu = E[U]$.

The variance is $\sigma^2 = E[(U - \mu)^2]$.

The standard deviation is $\sigma$.

It is always true that $\sigma^2 = E[U^2] - \mu^2$

For the exponential distribution

$$\sigma^2 = \frac{2}{a^2} - \frac{1}{a^2} = \frac{1}{a^2}$$

For an exponential distribution, $\sigma = \mu$. 
The $k^{th}$ moment of a random variable $U$ is defined as $E[U^k]$, and can be computed by the following formulas for the discrete and continuous cases, respectively.

$$E[U^k] = \sum_{i=1}^{L} u_i^k P[U = u_i]$$

$$E[U^k] = \int_{-\infty}^{\infty} u^k f_U(u) du$$
Example: Rayleigh Distribution

\[ f_R(r) = \frac{r}{b} e^{-r^2/2b} \quad (r \geq 0) \]

\[ \mu_R = E[R] = \int_0^\infty \frac{r^2}{b} e^{-r^2/2b} dr = \sqrt{\frac{\pi b}{2}} \]

\[ E[R^2] = \int_0^\infty \frac{r^3}{b} e^{-r^2/2b} dr = 2b \]

\[ \text{var}(R) = \sigma_R^2 = 2b - \frac{\pi b}{2} \approx 0.43b \]

The standard deviation is \( \sigma_R \approx 0.655\sqrt{b} \). By a little calculus you can show that the peak of the distribution occurs at \( r = \sqrt{b} \), so that the standard deviation is about 65% of the location of the peak. As \( b \) is increased the curve shifts to the right and broadens in proportion to \( \sqrt{b} \).
Multivariate Functions

Let $U = [U_1, U_2, \ldots, U_n]$ where $U_i, i = 1, \ldots, n$ are random variables.

Let $V = g(U)$. Then $V$ is a function of several random variables.

$$E[V] = \int \int \cdots \int g(u_1, u_2, \ldots, u_n) f_U(u_1, u_2, \ldots, u_n) du_1 du_2 \ldots du_n$$
Dartboard Example

Consider the problem of throwing darts at a target. Let $X$ and $Y$ be random variables that represent the horizontal and vertical offset of the dart from the center of the target. Note that the center is at $(0,0)$ so that the variables can have both negative and positive values. Suppose that $X$ and $Y$ are statistically independent normal random variables, with

$$f_X(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left[ -\frac{x^2}{2\sigma^2} \right]$$

$$f_Y(y) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left[ -\frac{y^2}{2\sigma^2} \right]$$
Dartboard Example

Because the horizontal and vertical offsets are statistically independent, the joint distribution is

\[ f_{X,Y}(x, y) = f_X(x)f_Y(y) = \frac{1}{\sigma^22\pi}\exp\left[-\frac{x^2 + y^2}{2\sigma^2}\right] \]

Find the distance the dart hits from the center. Let \( R = \sqrt{X^2 + Y^2} \)

Substitute \( r^2 = x^2 + y^2 \) in the probability density function.

The volume above a ring of thickness \( dr \) and mean radius \( r \) is \( 2\pi rf_{X,Y}(r)dr \). This must be the same value as achieved under the distribution of \( f_R(r) \) over the interval \( dr \). Therefore,

\[ f_R(r)dr = 2\pi rf_{X,Y}(r)dr \]
Dartboard Example

\[
f_R(r) = \frac{r}{\sigma^2} \exp \left( -\frac{r^2}{2\sigma^2} \right)
\]

Replace \(\sigma^2\) with a parameter \(b\). The result is

\[
f_R(r) = \frac{r}{b} \exp \left( -\frac{r^2}{2b} \right)
\]

This is the Rayleigh probability density or Rayleigh distribution with parameter \(b\).

What is the PDF on the angle \(\theta\)?
Characteristic Functions
Characteristic Function

The characteristic function of a random variable $X$ is the expected value of the function $e^{juX}$ where $j = \sqrt{-1}$.

**Continuous Distribution**

$$M_X(ju) = E\left[e^{juX}\right] = \int_{-\infty}^{\infty} f_X(x)e^{jux}dx$$

**Discrete Distribution**

$$M_X(ju) = E\left[e^{juX}\right] = \sum_{i} f_X(x_i)e^{jux_i}$$

The similarity of the characteristic function and the Fourier transform is obvious.

A major use of the cf is in generating moments of random variables.
Characteristic Function

Since $f_X(x)$ is nonnegative and $e^{jux}$ has unit magnitude for all values of $x$ and $u$,

$$\left| \int_{-\infty}^{\infty} f_X(x)e^{jux} \, dx \right| \leq \int_{-\infty}^{\infty} |f_X(x)e^{jux}| \, dx = \int_{-\infty}^{\infty} f_X(x) \, dx = 1$$

Hence, the characteristic function always exists and

$$|M_X(ju)| \leq M_X(0) = 1$$

If the cf is known, then one can compute the pdf by

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} M_X(ju)e^{-jux} \, du$$
Moment Generation

The function $e^{juX}$ can be expanded in a Taylor series:

$$e^{juX} = \sum_{n=0}^{\infty} \frac{(juX)^n}{n!} = \sum_{n=0}^{\infty} \frac{(ju)^n}{n!}X^n$$

Hence,

$$M_X(ju) = \sum_{n=0}^{\infty} \frac{(ju)^n}{n!}E[X^n] = \sum_{n=0}^{\infty} \frac{(ju)^n}{n!}m_n$$

where $m_n$ is the $n^{th}$ moment of $X$ about the origin.
Moment Generation

If we expand the characteristic function in a Taylor series, then $m_n$ can be found from the coefficient of $u^n$ in the expansion. Specifically, if

$$M_X(ju) = \sum_{n=0}^{\infty} k_n u^n$$

then $m_n = n! k_n / j^n$. This can be seen by comparison with

$$M_X(ju) = \sum_{n=0}^{\infty} \frac{(ju)^n}{n!} m_n$$

The moment can also be found by taking derivatives.

$$E[X^m] = \frac{1}{j^m} \left. \frac{d^m M_X}{du^m} \right|_{u=0}$$
Example: Exponential PDF

Consider the exponential pdf given by

\[ f_X(x) = be^{-bx} \text{step}(x) \]

The characteristic function is

\[ M_X(u) = E[e^{ju}] \]

\[ = \int_0^{\infty} be^{-(b-ju)x} dx \]

\[ = \frac{b}{b-ju} = \frac{1}{1 - \frac{ju}{b}} \]
Within the circle $|u| < b$ we can expand the term on the right in a power series by using (for $|t| < 1$)

\[
\frac{1}{1-t} = 1 + t + t^2 + t^3 + \cdots
\]

Then the expansion, with $t = ju/b$, is

\[
M_X(u) = \sum_{n=0}^{\infty} \left( \frac{ju}{b} \right)^n
\]

In this case $k_n = (j/b)^n$ so that

\[
m_n = E[X^n] = \frac{n!}{j^n k_n} = \frac{n!}{b^n}
\]
Example: Binomial Distribution

The probability of $k$ successes in $n$ trials is governed by the binomial probability distribution

$$b(n, k, p) = \binom{n}{k} p^k (1 - p)^{n-k}$$

The cf is

$$M_X(ju) = \sum_{k=0}^{n} b(n, k, p) e^{juk} = \sum_{k=0}^{n} \binom{n}{k} p^k (1 - p)^{n-k} e^{juk}$$

$$= \sum_{k=0}^{n} \binom{n}{k} (pe^{ju})^k (1 - p)^{n-k} = (pe^{ju} + 1 - p)^n$$

Here we have made use of the binomial expansion

$$(a + b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$
Binomial Distribution - continued

The mean value is found by differentiation.

\[ m_1 = E[X] = \frac{1}{j} \frac{dM}{du} \bigg|_{u=0} = npe^{ju}(pe^{ju} + 1 - p)^{n-1} \bigg|_{u=0} = np \]

Similarly,

\[ m_2 = E[X^2] = \frac{1}{j^2} \frac{d^2M}{du^2} \bigg|_{u=0} = n(n-1)p^2e^{ju}(pe^{ju} + 1 - p)^{n-2} + npe^{ju}(pe^{ju} + 1 - p)^{n-1} \bigg|_{u=0} = n(n-1)p^2 + np \]

The variance is easily found by

\[ \sigma^2 = \text{var}[X] = E[X^2] - E^2[X] = n(n-1)p^2 + np - n^2p^2 = np(1 - p) \]