10.1: Functions of a Random Variable

Introduction

Frequently, we observe a value of some random variable, but are really interested in a value derived from this by a function rule. If X is a random variable and g is a reasonable function (technically, a *Borel function*), then $Z = g(X)$ is a new random variable which has the value $g(t)$ for any ω such that $X(\omega) = t$. Thus $Z(\omega) = g(X(\omega))$.

The problem; an approach

We consider, first, functions of a single random variable. A wide variety of functions are utilized in practice.

Example 10.1.1: A quality control problem

In a quality control check on a production line for ball bearings it may be easier to weigh the balls than measure the diameters. If we can assume true spherical shape and w is the weight, then diameter is $kw^{1/3}$, where k is a factor depending upon the formula for the volume of a sphere, the units of measurement, and the density of the steel. Thus, if X is the weight of the sampled ball, the desired random variable is $D = kX^{1/3}$.

Example 10.1.2: Price breaks

The cultural committee of a student organization has arranged a special deal for tickets to a concert. The agreement is that the organization will purchase ten tickets at \$20 each (regardless of the number of individual buyers). Additional tickets are available according to the following schedule:

- 11-20, \$18 each
- 21-30, \$16 each
- 31-50, \$15 each
- 51-100, \$13 each

If the number of purchasers is a random variable X, the total cost (in dollars) is a random quantity $Z = g(X)$ described by

$$g(X) = 200 + 18I_{M1}(X)(X-10) + (16-18)I_{M2}(X)(X-20)$$

$$+(15-16)I_{M3}(X)(X-30) + (13-15)I_{M4}(X)(X-50)$$

$$\text{where } M1 = [10, \infty),\ M2 = [20, \infty),\ M3 = [30, \infty),\ M4 = [50, \infty)$$

The function rule is more complicated than in Example 10.1.1, but the essential problem is the same.

The problem

If X is a random variable, then $Z = g(X)$ is a new random variable. Suppose we have the distribution for X. How can we determine $P(Z \in M)$, the probability Z takes a value in the set M?

An approach to a solution

We consider two equivalent approaches

To find $P(X \in M)$.

1. *Mapping approach*. Simply find the amount of probability mass mapped into the set M by the random variable X.
   - In the absolutely continuous case, calculate $\int_M f_X$.
   - In the discrete case, identify those values t_i of X which are in the set M and add the associated probabilities.
2. *Discrete alternative*. Consider each value t_i of X. Select those which meet the defining conditions for M and add the associated probabilities. This is the approach we use in the MATLAB calculations. Note that it is not necessary to describe geometrically the set M; merely use the defining conditions.

To find $P(g(X) \in M)$.

1. *Mapping approach*. Determine the set N of all those t which are mapped into M by the function g. Now if $X(\omega) \in N$, then $g(X(\omega)) \in M$, and if $g(X(\omega)) \in M$, then $X(\omega) \in N$. Hence

$$\{\omega : g(X(\omega)) \in M\} = \{\omega : X(\omega) \in N\}$$

Since these are the same event, they must have the same probability. Once N is identified, determine $P(X \in N)$ in the usual manner (see part a, above).

- *Discrete alternative*. For each possible value t_i of X, determine whether $g(t_i)$ meets the defining condition for M. Select those t_i which do and add the associated probabilities.

— □

Remark. The set N in the mapping approach is called the *inverse image* $N = g^{-1}(M)$

Example 10.1.3: A discrete example

Suppose X has values -2, 0, 1, 3, 6, with respective probabilities 0.2, 0.1, 0.2, 0.3 0.2.

Consider $Z = g(X) = (X+1)(X-4)$. Determine $P(Z > 0)$.

Solution

First solution. The mapping approach

$g(t) = (t+1)(t-4)$. $N = \{t : g(t) > 0\}$ is the set of points to the left of –1 or to the right of 4. The X-values –2 and 6 lie in this set. Hence

$$P(g(X) > 0) = P(X = -2) + P(X = 6) = 0.2 + 0.2 = 0.4$$

Second solution. The discrete alternative

X =	-2	0	1	3	6
P X =	0.2	0.1	0.2	0.3	0.2
Z =	6	-4	-6	-4	14
Z > 0	1	0	0	0	1

Picking out and adding the indicated probabilities, we have

$$P(Z > 0) = 0.2 + 0.2 = 0.4$$

In this case (and often for "hand calculations") the mapping approach requires less calculation. However, for MATLAB calculations (as we show below), the discrete alternative is more readily implemented.

Example 10.1.4 . An absolutely continuous example

Suppose X ~ uniform [–3,7]. Then $f_X(t) = 0.1$, $-3 \le t \le 7$ (and zero elsewhere). Let

$$Z = g(X) = (X+1)(X-4)$$

Determine $P(Z > 0)$.

Solution

First we determine $N = \{t : g(t) > 0\}$. As in Example 10.1.3, $g(t) = (t+1)(t-4) > 0$ for $t < -1$ or $t > 4)$. Because of the uniform distribution, the integral of the density over any subinterval of $\{X, Y\}$ is 0.1 times the length of that subinterval. Thus, the desired probability is

$$P(g(X) > 0) = 0.1[(-1-(-3)) + (7-4)] = 0.5$$

We consider, next, some important examples.

Example 10.1.5: The normal distribution and standardized normal distribution

To show that if $X \sim N(\mu, \sigma^2)$ then

$$Z = g(X) = \frac{X - \mu}{\sigma} \; N(0,1)$$

VERIFICATION

We wish to show the denity function for Z is

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$$

Now

$$g(t) = \frac{t-\mu}{\sigma} \leq v \text{ iff } t \leq \sigma v + \mu$$

Hence, for given $M = (-\infty, v]$ the inverse image is $N = (-\infty, \sigma v + \mu]$, so that

$$F_Z(v) = P(Z \leq v) = P(Z \in M) = P(X \in N) = P(X \leq \sigma v + \mu) = F_X(\sigma v + \mu)$$

Since the density is the derivative of the distribution function,

$$f_Z(v) = F_Z'(v) = F_X'(v) = F_X'(\sigma v + \mu)\sigma = \sigma f_X(\sigma v + \mu)$$

Thus

$$f_Z(v) = \frac{\sigma}{\sigma\sqrt{2\pi}} \exp[-\frac{1}{2}(\frac{\sigma v + \mu - \mu}{\sigma})^2] = \frac{1}{\sqrt{2\pi}} e^{-v^2/2} = \varphi(v)$$

We conclude that $Z \sim N(0, 1)$.

Example 10.1.1

Suppose X has distribution function F_X. If it is absolutely continuous, the corresponding density is f_X. Consider $Z = aX + b$. Here $g(t) = at + b$, an affine function (linear plus a constant). Determine the distribution function for Z (and the density in the absolutely continuous case).

Solution

$F_Z(v) = P(Z \leq v) = P(aX + b \leq v)$

There are two cases

- $a > 0$:

$$F_Z(v) = P(X \leq \frac{v-b}{a}) = F_X(\frac{v-b}{a})$$

- $a < 0$

$$F_Z(v) = P(X \geq \frac{v-b}{a}) = P(X > \frac{v-b}{a}) + P(X = \frac{v-b}{a})$$

So that

$$F_Z(v) = 1 - F_X(\frac{v-b}{a}) + P(X = \frac{v-b}{a})$$

For the absolutely continuous case, $P(X = \frac{v-b}{a}) = 0$, and by differentiation

- for $a > 0$ $f_Z(v) = \frac{1}{a} f_X(\frac{v-b}{a})$
- for $a < 0$ $f_Z(v) = -\frac{1}{a} f_X(\frac{v-b}{a})$

Since for $a < 0$, $-a = |a|$, the two cases may be combined into one formula.

$$f_Z(v) = \frac{1}{|a|} f_X(\frac{v-b}{a})$$

Example 10.1.7: Completion of normal and standardized normal relationship

Suppose $Z \sim N(0,1)$. show that $X = \sigma Z + \mu \ (\sigma > 0)$ is $N(\mu, \sigma^2)$.

VERIFICATION

Use of the result of Example 10.1.6 on affine functions shows that

$$f_X(t) = \frac{1}{\sigma}\varphi(\frac{t-\mu}{\sigma}) = \frac{1}{\sigma\sqrt{2\pi}}\exp[-\frac{1}{2}(\frac{t-\mu}{\sigma})^2]$$

Example 10.1.8: Fractional power of a nonnegative random variable

Suppose $X \ge 0$ and $Z = g(X) = X^{1/a}$ for $a > 1$. Since for $t \ge 0$, $t^{1/a}$ is increasing, we have $0 \le t^{1/a} \le v$ iff $0 \le t \le v^a$. Thus

$$F_Z(v) = P(Z \le v) = P(X \le v^a) = F_X(v^a)$$

In the absolutely continuous case

$$f_Z(v) = F_Z^{'}(v) = f_X(v^a)av^{a-1}$$

Example 10.1.9: Fractional power of an exponentially distributed random variable

Suppose $X \sim$ exponential (λ). Then $Z = X^{1/a} \sim$ Weibull $(a, \lambda, 0)$.

According to the result of Example 10.1.8,

$$F_Z(t) = F_X(t^a) = 1 - e^{-\lambda t^a}$$

which is the distribution function for $Z \sim$ Weibull $(a, \lambda, 0)$.

Example 10.1.10: A simple approximation as a function of X

If X is a random variable, a simple function approximation may be constructed (see Distribution Approximations). We limit our discussion to the bounded case, in which the range of X is limited to a bounded interval $I = [a, b]$. Suppose I is partitioned into n subintervals by points t_i, $1 \le i \le n-1$, with $a = t_0$ and $b = t_n$. Let $M_i = [t_{i-1}, t_i)$ be the ith subinterval, $1 \le i \le n-1$ and $M_n = [t_{n-1}, t_n]$. Let $E_i = X^{-1}(M_i)$ be the set of points mapped into M_i by X. Then the E_i form a partition of the basic space Ω. For the given subdivision, we form a simple random variable X_s as follows. In each subinterval, pick a point $s_i, t_{i-1} \le s_i < t_i$. The simple random variable

$$X_s = \sum_{i=1}^n s_i I_{E_i}$$

approximates X to within the length of the largest subinterval M_i. Now $I_{E_i} = I_{M_i}(X)$, since $I_{E_i}(\omega) = 1$ iff $X(\omega) \in M_i$ iff $I_{M_i}(X(\omega)) = 1$. We may thus write

$$X_s = \sum_{i=1}^n s_i I_{M_i}(X), \text{ a function of } X$$

Use of MATLAB on simple random variables

For simple random variables, we use the discrete alternative approach, since this may be implemented easily with MATLAB. Suppose the distribution for X is expressed in the row vectors X and PX.

- We perform *array operations* on vector X to obtain

$$G = [g(t_1) g(t_2) \cdots g(t_n)]$$

- We use *relational* and *logical* operations on G to obtain a matrix M which has ones for those t_i (values of X) such that $g(t_i)$ satisfies the desired condition (and zeros elsewhere).
- The zero-one matrix M is used to select the the corresponding $p_i = P(X = t_i)$ and sum them by the taking the dot product of M and PX.

Example 10.1.11: Basic calculations for a function of a simple random variable

```
X = -5:10;                     % Values of X
PX = ibinom(15,0.6,0:15);      % Probabilities for X
G = (X + 6).*(X - 1).*(X - 8); % Array operations on X matrix to get G = g(X)
M = (G > - 100)&(G < 130);     % Relational and logical operations on G
PM = M*PX'                     % Sum of probabilities for selected values
PM =  0.4800
disp([X;G;M;PX]')              % Display of various matrices (as columns)
   -5.0000   78.0000    1.0000    0.0000
   -4.0000  120.0000    1.0000    0.0000
   -3.0000  132.0000         0    0.0003
   -2.0000  120.0000    1.0000    0.0016
   -1.0000   90.0000    1.0000    0.0074
         0   48.0000    1.0000    0.0245
    1.0000         0    1.0000    0.0612
    2.0000  -48.0000    1.0000    0.1181
    3.0000  -90.0000    1.0000    0.1771
    4.0000 -120.0000         0    0.2066
    5.0000 -132.0000         0    0.1859
    6.0000 -120.0000         0    0.1268
    7.0000  -78.0000    1.0000    0.0634
    8.0000         0    1.0000    0.0219
    9.0000  120.0000    1.0000    0.0047
   10.0000  288.0000         0    0.0005
[Z,PZ] = csort(G,PX);          % Sorting and consolidating to obtain
disp([Z;PZ]')                  % the distribution for Z = g(X)
 -132.0000    0.1859
 -120.0000    0.3334
  -90.0000    0.1771
  -78.0000    0.0634
  -48.0000    0.1181
         0    0.0832
   48.0000    0.0245
   78.0000    0.0000
   90.0000    0.0074
  120.0000    0.0064
  132.0000    0.0003
  288.0000    0.0005
P1 = (G<-120)*PX '           % Further calculation using G, PX
P1 =  0.1859
p1 = (Z<-120)*PZ'            % Alternate using Z, PZ
p1 =  0.1859
```

Example 10.1.12

$X = 10I_A + 18I_B + 10I_C$ with $\{A, B, C\}$ independent and $P =$ [0.60.30.5].

We calculate the distribution for X, then determine the distribution for

$$Z = X^{1/2} - X + 50$$

```
c = [10 18 10 0];
pm = minprob(0.1*[6 3 5]);
canonic
 Enter row vector of coefficients  c
 Enter row vector of minterm probabilities  pm
Use row matrices X and PX for calculations
Call for XDBN to view the distribution
disp(XDBN)
         0    0.1400
   10.0000    0.3500
   18.0000    0.0600
   20.0000    0.2100
   28.0000    0.1500
   38.0000    0.0900
G = sqrt(X) - X + 50;       % Formation of G matrix
[Z,PZ] = csort(G,PX);       % Sorts distinct values of g(X)
disp([Z;PZ]')               % consolidates probabilities
   18.1644    0.0900
   27.2915    0.1500
   34.4721    0.2100
   36.2426    0.0600
   43.1623    0.3500
   50.0000    0.1400
M = (Z < 20)|(Z >= 40)      % Direct use of Z distribution
M =    1     0     0     0     1     1
PZM = M*PZ'
PZM =  0.5800
```

Remark. Note that with the m-function csort, we may name the output as desired.

Example 10.1.13: Continuation of example 10.1.12, above.

```
H = 2*X.^2 - 3*X + 1;
[W,PW] = csort(H,PX)
W  =     1      171     595     741    1485    2775
PW =  0.1400  0.3500  0.0600  0.2100  0.1500  0.0900
```

Example 10.1.14: A discrete approximation

Suppose X has density function $f_X(t) = \frac{1}{2}(3t^2 + 2t)$ for $0 \le t \le 1$. Then $F_X(t) = \frac{1}{2}(t^3 + t^2)$. Let $Z = X^{1/2}$. We may use the approximation m-procedure tappr to obtain an approximate discrete distribution. Then we work with the approximating random variable as a simple random variable. Suppose we want $P(Z \le 0.8)$. Now $Z \le 0.8$ iff $X \le 0.8^2 = 0.64$. The desired probability may be calculated to be

$$P(Z \le 0.8) = F_X(0.64) = (0.64^3 + 0.64^2)/2 = 0.3359$$

Using the approximation procedure, we have

```
tappr
Enter matrix [a b] of x-range endpoints  [0 1]
Enter number of x approximation points  200
Enter density as a function of t  (3*t.^2 + 2*t)/2
Use row matrices X and PX as in the simple case
G = X.^(1/2);
M = G <= 0.8;
PM = M*PX'
PM =   0.3359       % Agrees quite closely with the theoretical
```