

Random Variable Algebra

MA 223

Kurt Bryan

Some Notation

If X is a normal random variable with mean μ and variance σ^2 then we write $X = N(\mu, \sigma^2)$. Thus $X = N(0, 1)$ means X is a standard normal random variable.

Given a random variable X you'll often see the notation $E(X)$ for the mean μ of X . The E is for "expected value", which is sort of what the mean is supposed to quantify. We write $V(X)$ for the variance σ^2 of X . It doesn't matter whether X is continuous or discrete. For a discrete random variable X with density function f we then have

$$\mu = E(X) = \sum_x xf(x), \quad \sigma^2 = V(X) = \sum_x (x - \mu)^2 f(x),$$

while for a continuous random variable X with density function f we have

$$\mu = E(X) = \int_{-\infty}^{\infty} xf(x) dx, \quad \sigma^2 = V(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx.$$

Recall also that for a continuous random variable X with density function f the function

$$F(x) = \int_{-\infty}^x f(u) du$$

is called the "cumulative distribution function" for X . It has the interpretation that $F(x) = P(X < x)$, that is, $F(x)$ is the probability that the $X < x$. From the Fundamental Theorem of Calculus we have $F'(x) = f(x)$.

Linear Transformation of a Random Variable

For the rest of this handout we'll be concerned primarily with continuous random variables, with a few exceptions, but most of the ideas work for discrete variables too.

Suppose we construct a new random variable, $Y = aX$ where a is some constant; equivalently, $X = Y/a$. We want to compute the distribution and density functions for Y . Let's assume for simplicity that $a > 0$.

Let G denote the cumulative distribution for Y and g the density function. Divide both sides of $Y \leq y$ by a to get $Y/a \leq y/a$, or simply $X \leq y/a$ (since $a > 0$ the inequality won't flip). The statements $Y \leq y$ and $X \leq y/a$ are thus completely equivalent, and so

$$P(Y \leq y) = P(X \leq y/a).$$

But $P(X \leq y/a) = F(y/a)$ and $P(Y \leq y) = G(y)$, so that

$$G(y) = F(y/a). \tag{1}$$

This expresses the cumulative distribution function for Y in terms of the cumulative distribution function of X .

If you differentiate both sides of (1) with respect to y (and remember $F' = f$ and $G' = g$) you obtain

$$g(y) = \frac{1}{a} f\left(\frac{y}{a}\right), \quad (2)$$

which expresses the density function for Y in terms of the density function for X . Geometrically, the graph of the density function g is just the graph of f stretched horizontally by a factor of a and compressed vertically by a factor $1/a$ (so area under it remains 1).

Example 1: Suppose that X is uniform on the interval $[0, 1]$, so $f(x) = 1$ for $0 \leq x \leq 1$, and f is zero outside this interval. Let $Y = 2X$. From equation (2) we get $g(y) = \frac{1}{2}f(y/2)$, or simply $g(y) = 1/2$ for $0 \leq y \leq 2$ and g equal to zero outside this interval.

Example 2: Suppose that $X = N(\mu, \sigma^2)$ and $Y = aX$. The density function for X is $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$. From equation (2) the density function for Y is

$$\begin{aligned} g(y) &= \frac{1}{a} f\left(\frac{y}{a}\right) \\ &= \frac{1}{a\sigma\sqrt{2\pi}} e^{-\frac{(y/a-\mu)^2}{2\sigma^2}} \\ &= \frac{1}{a\sigma\sqrt{2\pi}} e^{-\frac{(y-a\mu)^2}{2a^2\sigma^2}} \end{aligned}$$

The latter is exactly the density function for a normal variable with mean $a\mu$ and variance $a^2\sigma^2$. That is, $Y = N(a\mu, a^2\sigma^2)$.

Mean and Variance of the Transformed Variable

Consider the mean and variance of $Y = aX$ for an arbitrary random variable X with density $f(x)$. We have

$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} yg(y) dy = \frac{1}{a} \int_{-\infty}^{\infty} yf\left(\frac{y}{a}\right) dy, \\ V(Y) &= \frac{1}{a} \int_{-\infty}^{\infty} (y - E(Y))^2 g(y) dy = \frac{1}{a} \int_{-\infty}^{\infty} (y - E(Y))^2 f\left(\frac{y}{a}\right) dy. \end{aligned}$$

You can evaluate the integral for $E(Y)$ on the right above by making the substitution $x = y/a$ (so $y = ax$, $dy = a dx$, limits still $-\infty$ to ∞) to find

$$E(Y) = \frac{1}{a} \int_{-\infty}^{\infty} axf(x) a dx = a \int_{-\infty}^{\infty} xf(x) dx = aE(X).$$

If you jam $E(Y) = aE(X)$ into the integral for $V(Y)$ and do a similar change of variable you find that

$$V(Y) = \frac{1}{a} \int_{-\infty}^{\infty} (ax - aE(X))^2 f(x) a dx = a^2 \int_{-\infty}^{\infty} (x - E(X))^2 f(x) dx = a^2V(X).$$

In summary, multiplying a random variable by a constant a multiplies the mean by a , but SQUARES the variance (or multiplies the standard deviation by a).

More General Transformations (Optional)

Consider a more general transformation of a random variable, say let $Y = \phi(X)$ where ϕ is some nonlinear function. We will require, however, that ϕ satisfy the condition that $\phi(p) < \phi(q)$ whenever $p < q$ (i.e, ϕ is a strictly increasing function, so stuff like $\phi(x) = \sin(x)$ is not admissible). You no doubt recall from Calculus that this means ϕ is invertible on its range. In this case it's easy to see that $Y \leq y$ exactly when $\phi^{-1}(Y) \leq \phi^{-1}(y)$, i.e, $X \leq \phi^{-1}(y)$. Thus

$$P(Y \leq y) = P(X \leq \phi^{-1}(y))$$

or

$$G(y) = F(\phi^{-1}(y)).$$

That's the relation between the distribution functions. Differentiate to get a relation between the density functions,

$$g(y) = \frac{d(\phi^{-1}(y))}{dy} f(\phi^{-1}(y)).$$

Example: Suppose $X = N(0, 1)$ and $Y = e^X$. What does the density function of Y look like? Here we have $\phi(x) = e^x$, so $\phi^{-1}(x) = \ln(x)$ and $\frac{d(\phi^{-1}(x))}{dx} = 1/x$. The distribution of Y is given by

$$g(y) = \frac{1}{y\sqrt{2\pi}} e^{-(\ln(y))^2/2}$$

for $y > 0$ (since the sample space of Y clearly consists of only the positive reals).

Exercise 1

- Let X be a random variable with density function $f(x)$ and n a positive integer. The integral

$$M_n(X) = \int_{-\infty}^{\infty} x^n f(x) dx$$

is called the *n*th moment of X (so $M_1(X)$ is just the mean). Let $Y = aX$ for some positive constant a . How are the moments $M_n(Y)$ and $M_n(X)$ related?

- Let X be a uniform variable for $1 < x < 2$ and let $Y = e^X$. What is the density function for Y ? Compute it theoretically, then do a Minitab simulation and histogram, and compare.

Linear Combinations of Independent Random Variables—Discrete Case

Let X and Y be discrete random variables with density functions f and g . For simplicity let's suppose that X takes values $0, 1, 2, \dots, m$ and Y takes values $0, 1, 2, \dots, n$. The bivariate pair (X, Y) is a random variable in its own right and has some joint density function $h(x, y)$.

The interpretation is that $h(x, y)$ is the probability that $X = x$ AND $Y = y$ simultaneously. We say that X and Y are independent if

$$h(x, y) = f(x)g(y). \tag{3}$$

Note the similarity to the definition of independent events—the probability of both occurring simultaneously is the product of the individual probabilities. A similar definition is used if X and Y are continuous random variables. We say that X and Y are independent if the joint density function obeys equation (3).

Now let X and Y be independent variables with density functions f and g , respectively and define a new random variable $U = X + Y$. What is the density function $d(u)$ for U ? Let's first consider the case in which X and Y are both discrete as above, so that $U = X + Y$ can assume values $0, 1, 2, \dots, m + n$. What is the probability that $U = r$ for some fixed r ? This can happen if $X = 0$ and $Y = r$, or $X = 1$ and $Y = r - 1$, or $X = 2$ and $Y = r - 2$, etc., all the way up to $X = r$ and $Y = 0$. In each case we have $X = j$ and $Y = r - j$ for some j in the range $0 \leq j \leq r$, and since X and Y are independent the probability that $X = j$ and $Y = r - j$ simultaneously is $h(j, r - j) = f(j)g(r - j)$. Each of these (X, Y) pairs is disjoint, so the probability that $U = r$ is given by

$$\sum_{j=0}^{j=r} f(j)g(r - j).$$

Example: Suppose we roll two dice (six-sided). The dice are of course assumed to be independent, and the probability of any give face coming up are $1/6$. Just to keep it in the same framework as above, let's label the die faces $0, 1, 2, 3, 4, 5$ instead of 1 through 6 . What is the probability that the dice sum to 5 ? Let X denote the random variable corresponding to the “up” face of die 1 , Y the same for die 2 . Then $X + Y = 5$ can occur as any of the (X, Y) pairs $(0, 5), (1, 4), (2, 3), (3, 2), (4, 1),$ or $(5, 0)$. Each of these occurs with probability $(1/6)(1/6) = 1/36$, since the two dice are independent. The probability that $U = 5$ is

$$\sum_{j=0}^5 (1/6)(1/6) = 1/6.$$

Linear Combinations of Independent Random Variables—Continuous Case

Now let X_1 and X_2 be independent continuous random variables with density functions f and g , and define $Y = X_1 + X_2$. We want to compute the density function $d(y)$ for Y . Here's a very intuitive approach (which is easy enough to make rigorous—see the Appendix). Approach it like the discrete case above, by noting that for any fixed choice of y the value of $d(y)$ should be the probability that $X_1 + X_2 = y$. We can obtain this result if $X_1 = x$ and $X_2 = y - x$ for any x . Since X_1 and Y_2 are independent the probability of this occurring is $f(x)g(y - x)$. Each outcome $(X, Y) = (x, y - x)$ is disjoint, so the probability that $Y = y$ is given the by the “sum”

$$d(y) = \int_{-\infty}^{\infty} f(x)g(y - x) dx. \tag{4}$$

where I've assumed f and g are defined for all real values (set them to zero outside the sample space, if needed).

Example: Let X_1 be an exponential variable with density $f(x) = e^{-x}$ for $x \geq 0$, and let X_2 be exponential also, independent, with density $g(z) = 2e^{-2z}$ for $z \geq 0$. We define f and g to be zero for arguments less than 0. If $Y = X_1 + X_2$ then the density function for Y is

$$d(y) = \int_{-\infty}^{\infty} f(x)g(y-x) dx = \int_0^y f(x)g(y-x) dx = 2 \int_0^y e^{-x}e^{2(y-x)} dx.$$

Think about why the latter integral has limits 0 to y . Working the integral out yields

$$d(y) = 2e^{-y} - 2e^{-2y}.$$

Example: Let $X_1 = N(\mu_1, \sigma_1^2)$ (density $f_1(x) = \frac{e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}}{\sigma_1\sqrt{2\pi}}$) and $X_2 = N(\mu_2, \sigma_2^2)$ (density $f_2(x) = \frac{e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}}}{\sigma_2\sqrt{2\pi}}$). Let $Y = X_1 + X_2$, so that the density $d(y)$ for Y is given by

$$\begin{aligned} d(y) &= \int_{-\infty}^{\infty} f_1(x)f_2(y-x) dx \\ &= \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}}{\sigma_1\sqrt{2\pi}} \frac{e^{-\frac{(y-x-\mu_2)^2}{2\sigma_2^2}}}{\sigma_2\sqrt{2\pi}} dx \\ &= \frac{e^{-\frac{(y-(\mu_1+\mu_2))^2}{2(\sigma_1^2+\sigma_2^2)}}}{\sqrt{\sigma_1^2 + \sigma_2^2}\sqrt{2\pi}}. \end{aligned}$$

The transition to the last line is just mindless integration with the substitution $u = x - \mu_1$ (it also helps to let $\mu_1 + \mu_2 = \mu$). The last line in fact shows that $Y = N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$. We already knew what the mean and variance of Y would be, but the important point is that *the sum of two (or more) normals is again normal*.

It's not hard (conceptually) to compute the distribution of a sum $Y = X_1 + X_2 + \dots + X_n$ of independent random variables with density functions f_1, f_2, \dots, f_n : Just apply formula (4) to compute the density function for $X_1 + X_2$, then apply it again to get $X_1 + X_2 + X_3$, and so on. Of course the computations could be quite difficult, but at least the idea is straightforward enough.

Exercises 2

- Let X_1 and X_2 both be uniform on $(0, 1)$. Find the density function for $U = X_1 + X_2$.
- Repeat the last part for three uniform variables on $(0, 1)$, that is, compute the density for $U = X_1 + X_2 + X_3$.

Mean and Variance of Sums

Let X_1 and X_2 have density functions f_1 and f_2 . The density function for $Y = X_1 + X_2$ is given by equation (4). The expected value for Y is given by the integral

$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} y d(y) dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f_1(x) f_2(y-x) dx dy. \end{aligned}$$

Make a substitution $y = u + x$ (so $y - x = u$, $dy = du$, limits still $-\infty$ to ∞) to find

$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (u+x) f_1(x) f_2(u) dx du \\ &= \int_{-\infty}^{\infty} x f_1(x) \left(\int_{-\infty}^{\infty} f_2(u) du \right) dx + \int_{-\infty}^{\infty} u f_2(u) \left(\int_{-\infty}^{\infty} f_1(x) dx \right) du \\ &= \int_{-\infty}^{\infty} x f_1(x) dx + \int_{-\infty}^{\infty} u f_2(u) du \\ &= E(X_1) + E(X_2) \end{aligned}$$

where I've swapped orders of integration and used the fact that the integrals of f_1 and f_2 equal 1. So the expected value of the sum is the sum of the expected values! Not surprising, if you think about it.

The same thing turns out to be true for the variances. Let $\mu_1 = E(X_1)$ and $\mu_2 = E(X_2)$. Then the variance of Y is given by

$$\begin{aligned} V(Y) &= \int_{-\infty}^{\infty} (y - \mu_1 - \mu_2)^2 g(y) dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y - \mu_1 - \mu_2)^2 f_1(x) f_2(y-x) dx dy. \end{aligned}$$

Do the same change of variables, $y = u + x$, to find

$$\begin{aligned} V(Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ((x - \mu_1) + (u - \mu_2))^2 f_1(x) f_2(u) dx du \\ &= \int_{-\infty}^{\infty} (x - \mu_1)^2 f_1(x) \left(\int_{-\infty}^{\infty} f_2(u) du \right) dx + \left(\int_{-\infty}^{\infty} (x - \mu_1) f_1(x) dx \right) \left(\int_{-\infty}^{\infty} (u - \mu_2) f_2(u) du \right) \\ &\quad + \int_{-\infty}^{\infty} (u - \mu_2)^2 f_2(u) \left(\int_{-\infty}^{\infty} f_1(x) dx \right) du \\ &= \int_{-\infty}^{\infty} (x - \mu_1)^2 f_1(x) \left(\int_{-\infty}^{\infty} f_2(u) du \right) dx + \int_{-\infty}^{\infty} (u - \mu_2)^2 f_2(u) \left(\int_{-\infty}^{\infty} f_1(x) dx \right) du \\ &= V(X_1) + V(X_2). \end{aligned}$$

Think about why the cross-term $\left(\int_{-\infty}^{\infty} (x - \mu_1) f_1(x) dx \right) \left(\int_{-\infty}^{\infty} (u - \mu_2) f_2(u) du \right)$ in the second line is zero.

So variances add just like means if the random variables are independent.

Exercise 3:

- Suppose that X_1 is uniform on $(0, 5)$ and X_2 is normal with mean 3 and variance 4. Find the mean and variance of $Y = X_1 + X_2$ (but you don't need to find the density of Y !). Do a Minitab simulation with 1000 samples of X_1 and X_2 (add them to get Y), then use the `describe` command to confirm your result.

By iterating it's clear that the same results will hold for any number of random variables, e.g., $E(X_1 + X_2 + \dots + X_n) = E(X_1) + \dots + E(X_n)$. Also, since we know that $E(aX) = aE(X)$ and $V(aX) = a^2V(X)$, we can state a BIG result:

Theorem: If X_1, X_2, \dots, X_n are independent random variables with finite expected values $E(X_1), \dots, E(X_n)$ and variances $V(X_1), \dots, V(X_n)$, and $Y = a_1X_1 + \dots + a_nX_n$ for constants a_1, a_2, \dots, a_n then

$$\begin{aligned} E(Y) &= a_1E(X_1) + \dots + a_nE(X_n) \\ V(Y) &= a_1^2V(X_1) + \dots + a_n^2V(X_n) \end{aligned}$$

Exercise 4:

- Suppose X_1 and X_2 are normally distributed, X_1 with mean 3 and variance 8, X_2 also with mean 3 and variance 8. Compute the mean and variance of $X_1 - X_2$. Do a Minitab simulation to confirm your result.
- Suppose that X_1, X_2, \dots, X_n all have the same mean μ and variance σ^2 . Find the mean and variance of

$$Y = \frac{1}{n}(X_1 + X_2 + \dots + X_n)$$

in terms of μ, σ^2 , and n . What do $E(Y)$ and $V(Y)$ do as $n \rightarrow \infty$?

Appendix

Let $U = X + Y$, where X and Y are independent continuous random variables with density functions f and g , respectively. We want to compute the density function d for u . First note that

$$d(u) = \lim_{\Delta u \rightarrow 0^+} \frac{1}{\Delta u} \int_u^{u+\Delta u} d(z) dz \quad (5)$$

at least if $d(u)$ is continuous. For a fixed positive value of Δu we can interpret the integral on the right in (5) as $P(u < U < u + \Delta u)$, or $P(u < X + Y < u + \Delta u)$. The region $u < x + y < u + \Delta u$ in the xy -plane is a thin strip at a 45 degree angle (downward, as x increases); see Figure 1. If the pair (X, Y) has joint density function $h(x, y)$ then $P(u < X + Y < u + \Delta u)$ can be computed as the double integral of $h(x, y)$ over this region. The result is

$$P(u < X + Y < u + \Delta u) = \int_{-\infty}^{\infty} \int_{u-x}^{u-x+\Delta u} h(x, y) dy dx.$$

But since X and Y are independent $h(x, y) = f(x)g(y)$, so we find

$$P(u < X + Y < u + \Delta u) = \int_{-\infty}^{\infty} \int_{u-x}^{u-x+\Delta u} f(x)g(y) dy dx.$$

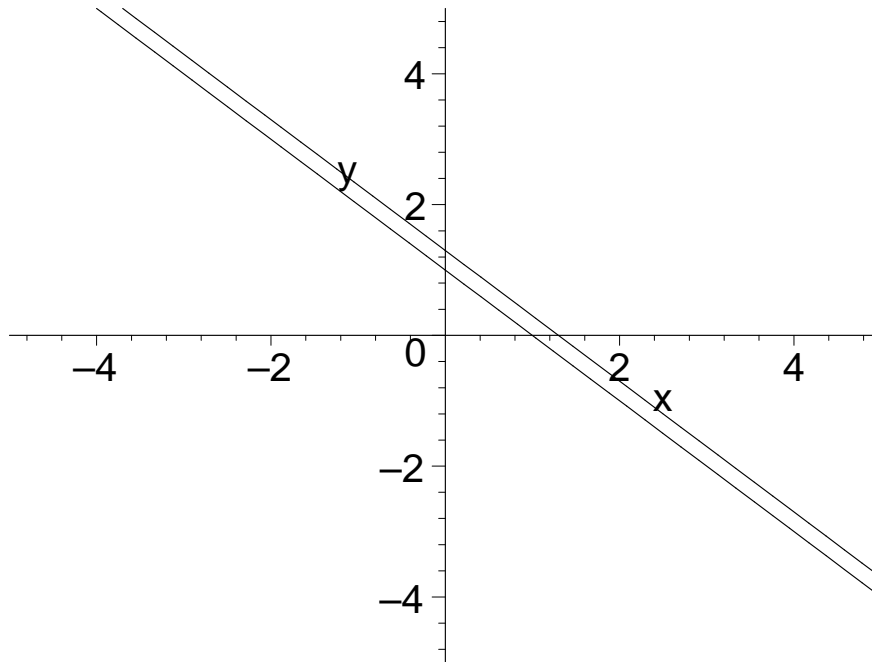


Figure 1: Region $u < x + y < u + \Delta u$ in xy -plane.

If we divide both sides above by Δu and let $\Delta u \rightarrow 0$ then the inside integral in y just approaches $g(u - x)$ and we obtain

$$\lim_{\Delta u \rightarrow 0^+} \frac{P(u < X + Y < u + \Delta u)}{\Delta u} = \lim_{\Delta u \rightarrow 0^+} \frac{1}{\Delta u} \int_{-\infty}^{\infty} f(x)g(u - x) dx.$$

The left side is (from equation (5)) the quantity $d(u)$.