Standardizing random variables

Linear combinations of random variables

Suppose X is a random variable (with well defined expected value and standard deviation) and b is a number. Then $X+b$ is another random variable, and

$$E(X+b) = E(x) + b$$
$$\text{st.dev.}(X+b) = \text{st.dev.}(X) \ .$$

If c is a number, then

$$E(cX) = cE(x)$$
$$\text{st.dev.}(cX) = |c|\text{st.dev.}(X) \ .$$

The standardization of a random variable

Suppose X is a random variable with mean μ and standard deviation $\sigma > 0$. Then the *standardization* of X is the random variable $Z = (X-\mu)/\sigma$. Then Z has mean zero and standard deviation 1.
(Changing X to $X-\mu$ changes the mean to zero and does not change the standard deviation. Then multiplying by $1/\sigma$ changes the standard deviation to 1, and doesn't change the mean, because $(1/\sigma)0 = 0$.)

If c and d are numbers, and c is nonzero, then X and $cX+d$ have the same standardization. Standardization gives us standard units for considering (for example) the shape the graph of a probability density function. If X records experimental measurements in feet, and Y records the experimental measurements in inches, and X and Y measure the same experiment in the lab, then their standardizations will be the same.

Even more important, standardization gives us a way to see the pattern of sums and averages.

For example, suppose $S_n = X_1 + \cdots + X_n$, and the X_i are i.i.d. with mean $\mu > 0$ and standard deviation $\sigma > 0$. How can we understand the pattern of S_n as n increases?

The expected value of S_n is $n\mu$. The outputs of S_n keeping getting larger. There doesn't seem to be any convergence.

If we replace S_n with $S_n - n\mu$, we at least see how the values of S_n are arranged around its mean. But the standard deviation of S_n is $\sqrt{n}\,\sigma$; if we just look at S_n, we just see the spread of those values going to infinity.

But if we standardize S_n, we get

$$Z_n \;=\; \frac{(S_n - n\mu)}{\sqrt{n}\,\sigma}\;.$$

We'll see with the Central Limit Theorem that for large n the distribution of Z_n is approximately $\mathcal{N}(0,1)$. And we can use that to get information about S_n, and also averages.

Standardization and a probability density function

Suppose f is a p.d.f. for a continuous random variable X, with a mean μ and nonzero standard deviation σ.

The graph of the p.d.f (call it g) for the r.v. $X - \sigma$ is the graph of f shifted μ units to the left.

The p.d.f. (call it h) of the standardization $Z = (X - \mu)/\sigma$ is obtained from g as follows. $h(x) = \sigma g(x/\sigma)$.

For example, suppose $\sigma = 3$. Then the graph is compressed horizontally by a factor of three and expanded vertically by a factor of 3.