

3.13: The Expected Value of a Function of Several Variables and the Central Limit Theorem

We can extend the idea of an expected value to a function of multiple random variables. Let U and V be distributions whose random variables are u and v , respectively. Let the probability density functions for these distributions be $df_u(u)/du$ and $df_v(v)/dv$. In general, these probability density functions are different functions; that is, U and V are different distributions. Let $g(u, v)$ be some function of these random variables. The probability that an observation made on U produces a value of u in the range $u^* < u < u^* + du$ is

$$P(u^* < u < u^* + du) = \frac{df_u(u^*)}{du} du$$

and the probability that an observation made on V produces a value of v in the range $v^* < v < v^* + dv$ is

$$P(v^* < v < v^* + dv) = \frac{df_v(v^*)}{dv} dv$$

The probability that making one observation on each of these distributions produces a value of u that lies in the range $u^* < u < u^* + du$ and a value of v that lies in the range $v^* < v < v^* + dv$ is

$$\frac{df_u(u^*)}{du} \frac{df_v(v^*)}{dv} du dv$$

In a straightforward generalization, we define the expected value of $g(u, v)$, $\langle g(u, v) \rangle$, as

$$\langle g(u, v) \rangle = \int_{v=-\infty}^{\infty} \int_{u=-\infty}^{\infty} g(u, v) \frac{df_u(u)}{du} \frac{df_v(v)}{dv} du dv$$

If $g(u, v)$ is a sum of functions of independent variables, $g(u, v) = h(u) + k(v)$, we have

$$\begin{aligned} \langle g(u, v) \rangle &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [h(u) + k(v)] \frac{df_u(u)}{du} \frac{df_v(v)}{dv} du dv = \int_{-\infty}^{\infty} h(u) \frac{df_u(u)}{du} du + \int_{-\infty}^{\infty} k(v) \frac{df_v(v)}{dv} dv \\ &= \langle h(u) \rangle + \langle k(v) \rangle \end{aligned}$$

If $g(u, v)$ is a product of independent functions, $g(u, v) = h(u) k(v)$, we have

$$\begin{aligned} \langle g(u, v) \rangle &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(u) k(v) \frac{df_u(u)}{du} \frac{df_v(v)}{dv} du dv = \int_{-\infty}^{\infty} h(u) \frac{df_u(u)}{du} du \times \int_{-\infty}^{\infty} k(v) \frac{df_v(v)}{dv} dv \\ &= \langle h(u) \rangle \langle k(v) \rangle \end{aligned}$$

We can extend these conclusions to functions of the random variables of any number of distributions. If u_i is the random variable of distribution U_i whose probability density function is $df_i(u_i)/du_i$, the expected value of

$$g(u_1, \dots, u_i, \dots, u_N) = h_1(u_1) + \dots + h_i(u_i) + \dots + h_N(u_N)$$

becomes

$$\langle g(u_1, \dots, u_i, \dots, u_N) \rangle = \sum_{i=1}^N \langle h_i(u_i) \rangle$$

and the expected value of

$$g(u_1, \dots, u_i, \dots, u_N) = h_1(u_1) \dots h_i(u_i) \dots h_N(u_N)$$

becomes

$$\langle g(u_1, \dots, u_i, \dots, u_N) \rangle = \prod_{i=1}^N \langle h_i(u_i) \rangle$$

We are particularly interested in expected values for repeated trials made on the same distribution. We consider distributions for which the outcome of one trial is independent of the outcome of any other trial. The probability density function is the same for

every trial, so we have $f(u) = f_1(u_1) = \dots = f_i(u_i) = \dots = f_N(u_N)$. Let the values obtained for the random variable in a series of trials on the same distribution be $u_1, \dots, u_i, \dots, u_N$. For each trial, we have

$$\langle h_i(u_i) \rangle = \int_{-\infty}^{\infty} h_i(u_i) \frac{df_i(u_i)}{du_i} du_i$$

If we consider the special case of repeated trials in which the functions $h_i(u_i)$ are all the same function, so that $h(u) = h_1(u_1) = \dots = h_i(u_i) = \dots = h_N(u_N)$, the expected value of

$$\begin{aligned} g(u_1, \dots, u_i, \dots, u_N) \\ = h_1(u_1) + \dots + h_i(u_i) + \dots + h_N(u_N) \end{aligned}$$

becomes

$$\langle g(u_1, \dots, u_i, \dots, u_N) \rangle = \sum_{i=1}^N \langle h_i(u_i) \rangle = N \langle h(u) \rangle$$

and the expected value of

$$g(u_1, \dots, u_i, \dots, u_N) = h_1(u_1) \dots h_i(u_i) \dots h_N(u_N)$$

becomes

$$\langle g(u_1, \dots, u_i, \dots, u_N) \rangle = \prod_{i=1}^N \langle h_i(u_i) \rangle = \langle h(u) \rangle^N$$

Now let us consider N independent trials on the same distribution and let $h_i(u_i) = h(u_i) = u_i$. Then, the expected value of

$$g(u_1, \dots, u_i, \dots, u_N) = h_1(u_1) + \dots + h_i(u_i) + \dots + h_N(u_N)$$

becomes

$$\langle u_1 + \dots + u_i + \dots + u_N \rangle = \sum_{i=1}^N \langle u_i \rangle = N \langle u \rangle = N\mu$$

By definition, the average of N repeated trials is

$\bar{u}_N = (u_1 + \dots + u_i + \dots + u_N)/N$, so that the expected value of the mean of a distribution of an average-of- N repeated trials is

$$\langle \bar{u}_N \rangle = \frac{\langle u_1 + \dots + u_i + \dots + u_N \rangle}{N} = \mu$$

This proves one element of the [central limit theorem](#): The mean of a distribution of averages-of- N values of a random variable drawn from a parent distribution is equal to the mean of the parent distribution.

The variance of these averages-of- N is

$$\begin{aligned} \sigma_N^2 &= \langle (\bar{u}_N - \mu)^2 \rangle = \left\langle \left[\left(\frac{1}{N} \sum_{i=1}^N u_i \right) - \mu \right]^2 \right\rangle = \left\langle \left[\left(\frac{1}{N} \sum_{i=1}^N u_i \right) - \frac{N\mu}{N} \right]^2 \right\rangle = \frac{1}{N^2} \left\langle \left[\left(\sum_{i=1}^N u_i \right) - N\mu \right]^2 \right\rangle \\ &= \frac{1}{N^2} \left\langle \left[\left(\sum_{i=1}^N (u_i - \mu) \right) \right]^2 \right\rangle = \frac{1}{N^2} \left\langle \sum_{i=1}^N (u_i - \mu)^2 \right\rangle + \frac{2}{N^2} \left\langle \sum_{i=1}^{N-1} (u_i - \mu) \sum_{j=i+1}^N (u_j - \mu) \right\rangle = \frac{1}{N^2} \\ &\quad \sum_{i=1}^N \langle (u_i - \mu)^2 \rangle + \frac{2}{N^2} \left\langle \sum_{i=1}^{N-1} (u_i - \mu) \right\rangle \left\langle \sum_{j=i+1}^N (u_j - \mu) \right\rangle \end{aligned}$$

Where the last term is zero, because

$$\left\langle \sum_{i=1}^{N-1} (u_i - \mu) \right\rangle = \sum_{i=1}^{N-1} \langle (u_i - \mu) \rangle$$

and

$$\langle (u_i - \mu) \rangle = 0$$

By definition, $\sigma^2 = \langle (u_i - \mu)^2 \rangle$, so that we have

$$\sigma_N^2 = \frac{N\sigma^2}{N^2} = \frac{\sigma^2}{N}$$

This proves a second element of the central limit theorem: The variance of an average of N values of a random variable drawn from a parent distribution is equal to the variance of the parent distribution divided by N .

This page titled [3.13: The Expected Value of a Function of Several Variables and the Central Limit Theorem](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Paul Ellgen](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.