

5.1: Location-Scale Families

General Theory

As usual, our starting point is a random experiment modeled by a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, so that Ω is the set of outcomes, \mathcal{F} the collection of events, and \mathbb{P} the probability measure on the sample space (Ω, \mathcal{F}) . In this section, we assume that we fixed random variable Z defined on the probability space, taking values in \mathbb{R} .

Definition

For $a \in \mathbb{R}$ and $b \in (0, \infty)$, let $X = a + bZ$. The two-parameter family of distributions associated with X is called the *location-scale family* associated with the given distribution of Z . Specifically, a is the *location parameter* and b the *scale parameter*.

Thus a *linear transformation*, with positive slope, of the underlying random variable Z creates a *location-scale family* for the underlying distribution. In the special case that $b = 1$, the one-parameter family is called the *location family* associated with the given distribution, and in the special case that $a = 0$, the one-parameter family is called the *scale family* associated with the given distribution. Scale transformations, as the name suggests, occur naturally when physical units are changed. For example, if a random variable represents the length of an object, then a change of units from meters to inches corresponds to a scale transformation. Location transformations often occur when the zero reference point is changed, in measuring distance or time, for example. Location-scale transformations can also occur with a change of physical units. For example, if a random variable represents the temperature of an object, then a change of units from Fahrenheit to Celsius corresponds to a location-scale transformation.

Distribution Functions

Our goal is to relate various functions that determine the distribution of $X = a + bZ$ to the corresponding functions for Z . First we consider the (cumulative) distribution function.

If Z has distribution function G then X has distribution function F given by

$$F(x) = G\left(\frac{x-a}{b}\right), \quad x \in \mathbb{R} \quad (5.1.1)$$

Proof

For $x \in \mathbb{R}$

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(a + bZ \leq x) = \mathbb{P}\left(Z \leq \frac{x-a}{b}\right) = G\left(\frac{x-a}{b}\right) \quad (5.1.2)$$

Next we consider the probability density function. The results are a bit different for discrete distributions and continuous distribution, not surprising since the density function has different meanings in these two cases.

If Z has a discrete distribution with probability density function g then X also has a discrete distribution, with probability density function f given by

$$f(x) = g\left(\frac{x-a}{b}\right), \quad x \in \mathbb{R} \quad (5.1.3)$$

Proof

Z takes values in a countable subset $S \subset \mathbb{R}$ and hence X takes values in $T = \{a + bz : z \in S\}$, which is also countable. Moreover

$$f(x) = \mathbb{P}(X = x) = \mathbb{P}\left(Z = \frac{x-a}{b}\right) = g\left(\frac{x-a}{b}\right), \quad x \in \mathbb{R} \quad (5.1.4)$$

If Z has a continuous distribution with probability density function g , then X also has a continuous distribution, with probability density function f given by

$$f(x) = \frac{1}{b} g\left(\frac{x-a}{b}\right), \quad x \in \mathbb{R} \quad (5.1.5)$$

1. For the location family associated with g , the graph of f is obtained by shifting the graph of g , a units to the right if $a > 0$ and $-a$ units to the left if $a < 0$.
2. For the scale family associated with g , if $b > 1$, the graph of f is obtained from the graph of g by stretching horizontally and compressing vertically, by a factor of b . If $0 < b < 1$, the graph of f is obtained from the graph of g by compressing horizontally and stretching vertically, by a factor of b .

Proof

First note that $\mathbb{P}(X = x) = \mathbb{P}\left(Z = \frac{x-a}{b}\right) = 0$, so X has a continuous distribution. Typically, Z takes values in an interval of \mathbb{R} and thus so does X . The formula for the density function follows by taking derivatives of the [distribution function](#) above, since $f = F'$ and $g = G'$.

If Z has a mode at z , then X has a mode at $x = a + bz$.

Proof

This follows from [density function](#) in the discrete case or the [density function](#) in the continuous case. If g has a maximum at z then f has a maximum at $x = a + bz$.

Next we relate the quantile functions of Z and X .

If G and F are the distribution functions of Z and X , respectively, then

1. $F^{-1}(p) = a + b G^{-1}(p)$ for $p \in (0, 1)$
2. If z is a quantile of order p for Z then $x = a + bz$ is a quantile of order p for X .

Proof

These results follow from the [distribution function](#) above.

Suppose now that Z has a continuous distribution on $[0, \infty)$, and that we think of Z as the failure time of a device (or the time of death of an organism). Let $X = bZ$ where $b \in [0, \infty)$, so that the distribution of X is the scale family associated with the distribution of Z . Then X also has a continuous distribution on $[0, \infty)$ and can also be thought of as the failure time of a device (perhaps in different units).

Let G^c and F^c denote the reliability functions of Z and X respectively, and let r and R denote the failure rate functions of Z and X , respectively. Then

1. $F^c(x) = G^c(x/b)$ for $x \in [0, \infty)$
2. $R(x) = \frac{1}{b} r\left(\frac{x}{b}\right)$ for $x \in [0, \infty)$

Proof

Recall that $G^c = 1 - G$, $F^c = 1 - F$, $r = g/\bar{G}$, and $R = f/\bar{F}$. Thus the results follow from the [distribution function](#) and the [density function](#) above.

Moments

The following theorem relates the mean, variance, and standard deviation of Z and X .

As before, suppose that $X = a + bZ$. Then

1. $\mathbb{E}(X) = a + b \mathbb{E}(Z)$
2. $\text{var}(X) = b^2 \text{var}(Z)$
3. $\text{sd}(X) = b \text{sd}(Z)$

Proof

These result follow immediately from basic properties of expected value and variance.

Recall that the *standard score* of a random variable is obtained by subtracting the mean and dividing by the standard deviation. The standard score is dimensionless (that is, has no physical units) and measures the distance from the mean to the random variable in standard deviations. Since location-scale families essentially correspond to a change of units, it's not surprising that the standard score is unchanged by a location-scale transformation.

The standard scores of X and Z are the same:

$$\frac{X - \mathbb{E}(X)}{\text{sd}(X)} = \frac{Z - \mathbb{E}(Z)}{\text{sd}(Z)} \quad (5.1.6)$$

Proof

From the [mean and variance](#) above:

$$\frac{X - \mathbb{E}(X)}{\text{sd}(X)} = \frac{a + bZ - [a + b\mathbb{E}(Z)]}{b\text{sd}(Z)} = \frac{Z - \mathbb{E}(Z)}{\text{sd}(Z)} \quad (5.1.7)$$

Recall that the skewness and kurtosis of a random variable are the third and fourth moments, respectively, of the standard score. Thus it follows from the previous result that skewness and kurtosis are unchanged by location-scale transformations: $\text{skew}(X) = \text{skew}(Z)$, $\text{kurt}(X) = \text{kurt}(Z)$.

We can represent the moments of X (about 0) to those of Z by means of the binomial theorem:

$$\mathbb{E}(X^n) = \sum_{k=0}^n \binom{n}{k} b^k a^{n-k} \mathbb{E}(Z^k), \quad n \in \mathbb{N} \quad (5.1.8)$$

Of course, the moments of X about the location parameter a have a simple representation in terms of the moments of Z about 0:

$$\mathbb{E}[(X - a)^n] = b^n \mathbb{E}(Z^n), \quad n \in \mathbb{N} \quad (5.1.9)$$

The following exercise relates the moment generating functions of Z and X .

If Z has moment generating function m then X has moment generating function M given by

$$M(t) = e^{at} m(bt) \quad (5.1.10)$$

Proof

$$M(t) = \mathbb{E}(e^{tX}) = \mathbb{E}[e^{t(a+bZ)}] = e^{ta} \mathbb{E}(e^{tbZ}) = e^{at} m(bt) \quad (5.1.11)$$

Type

As we noted earlier, two probability distributions that are related by a location-scale transformation can be thought of as governing the same underlying random quantity, but in different physical units. This relationship is important enough to deserve a name.

Suppose that P and Q are probability distributions on \mathbb{R} with distribution functions F and G , respectively. Then P and Q are of the same *type* if there exist constants $a \in \mathbb{R}$ and $b \in (0, \infty)$ such that

$$F(x) = G\left(\frac{x - a}{b}\right), \quad x \in \mathbb{R} \quad (5.1.12)$$

Being of the same type is an equivalence relation on the collection of probability distributions on \mathbb{R} . That is, if P , Q , and R are probability distribution on \mathbb{R} then

1. P is the same type as P (the *reflexive* property).
2. If P is the same type as Q then Q is the same type as P (the *symmetric* property).
3. If P is the same type as Q , and Q is the same type as R , then P is the same type as R (the *transitive* property).

Proof

Let F , G , and H denote the distribution functions of P , Q , and R respectively.

1. This is trivial, of course, since we can take $a = 0$ and $b = 1$.

2. Suppose there exists $a \in \mathbb{R}$ and $b \in (0, \infty)$ such that $F(x) = G\left(\frac{x-a}{b}\right)$ for $x \in \mathbb{R}$. Then

$$G(x) = F(a + bx) = F\left(\frac{x - (-a/b)}{1/b}\right) \text{ for } x \in \mathbb{R}.$$

3. Suppose there exists $a, c \in \mathbb{R}$ and $b, d \in (0, \infty)$ such that $F(x) = G\left(\frac{x-a}{b}\right)$ and $G(x) = H\left(\frac{x-c}{d}\right)$ for $x \in \mathbb{R}$. Then

$$F(x) = H\left(\frac{x - (a+bc)}{bd}\right) \text{ for } x \in \mathbb{R}.$$

So, the collection of probability distributions on \mathbb{R} is partitioned into mutually exclusive equivalence classes, where the distributions in each class are all of the same type.

Examples and Applications

Special Distributions

Many of the special parametric families of distributions studied in this chapter and elsewhere in this text are location and/or scale families.

The [arcsine distribution](#) is a location-scale family.

The [Cauchy distribution](#) is a location-scale family.

The [exponential distribution](#) is a scale family.

The [exponential-logarithmic distribution](#) is a scale family for each value of the shape parameter.

The [extreme value distribution](#) is a location-scale family.

The [gamma distribution](#) is a scale family for each value of the shape parameter.

The [Gompertz distribution](#) is a scale family for each value of the shape parameter.

The [half-normal distribution](#) is a scale family.

The [hyperbolic secant](#) distribution is a location-scale family.

The [Lévy distribution](#) is a location scale family.

The [logistic distribution](#) is a location-scale family.

The [log-logistic distribution](#) is a scale family for each value of the shape parameter.

The [Maxwell distribution](#) is a scale family.

The [normal distribution](#) is a location-scale family.

The [Pareto distribution](#) is a scale family for each value of the shape parameter.

The [Rayleigh distribution](#) is a scale family.

The [semicircle distribution](#) is a location-scale family.

The [triangle distribution](#) is a location-scale family for each value of the shape parameter.

The [uniform distribution](#) on an interval is a location-scale family.

The [U-power distribution](#) is a location-scale family for each value of the shape parameter.

The [Weibull distribution](#) is a scale family for each value of the shape parameter.

The [Wald distribution](#) is a scale family, although in the usual formulation, neither of the parameters is a scale parameter.

This page titled [5.1: Location-Scale Families](#) is shared under a [CC BY 2.0](#) license and was authored, remixed, and/or curated by [Kyle Siegrist](#) ([Random Services](#)) via [source content](#) that was edited to the style and standards of the LibreTexts platform.