Skew

Overview

Chebychev's theorem states that mean and median cannot differ by more than one standard deviation. This sets the coefficient of skew at plus/minus 3.

Tail to the right indicates positive skew where mode < median < **mean**
Tail to the left indicates positive skew where **mean** < median < mode

So for a simple continuous distribution with smooth tails, the position of the **mean** in the distribution indicates positive or negative

Note the measures of central tendency are in **alphabetical** order.

If **median** closer to Q1 lower quartile skew is likely to be positive
If **median** closer to Q3 upper quartile skew is likely to be negative

Measures of Skew

Parametric Measure *(taken **from** the actual **values** of the variable X)*

Fisher's moment of skew $\mu_3 = E [(X - \mu)^3 / \sigma]$

Non Parametric Measures

Traditional non parametric measure is S = (mean – median) / std. dev.

Pearson's 1st measure of skew = (mean – mode) / std. dev.

Pearson's 2nd measure of skew = 3 (mean – mode) / std. dev.

Quartile measure is { (Q3 – Q2) – (Q2 – Q1) } / { Q3 – Q1 }

Complications

For a unimodal distribution (a distribution with a single peak), negative skew commonly indicates that the tail is on the left side of the distribution, and positive skew indicates that the tail is on the right. In cases where one tail is long but the other tail is fat, skewness does not obey a simple rule. For example, a zero value in skewness means that the tails on both sides of the mean balance out overall; this is the case for a symmetric distribution but can also be true for an asymmetric distribution where one tail is long and thin, and the other is short but fat. Thus, the judgement on the symmetry of a given distribution by using only its skewness is risky; the distribution shape must be taken into account.

Skew

Overview

Chebychev's theorem states that mean and median cannot differ by more than one standard deviation. This sets the coefficient of skew at plus/minus 3.

Tail to the right indicates positive skew where mode < median < **mean**
Tail to the left indicates positive skew where **mean** < median < mode

So for a simple continuous distribution with smooth tails, the position of the **mean** in the distribution indicates positive or negative

Note the measures of central tendency are in **alphabetical** order.

If **median** closer to Q1 lower quartile skew is likely to be positive
If **median** closer to Q3 upper quartile skew is likely to be negative

Measures of Skew

Parametric Measure *(taken **from** the actual **values** of the variable X)*

Fisher's moment of skew $\mu_3 = E[(X - \mu)^3 / \sigma]$

Non Parametric Measures

Traditional non parametric measure is S = (mean – median) / std. dev.

Pearson's 1st measure of skew = (mean – mode) / std. dev.

Pearson's 2nd measure of skew = 3 (mean – mode) / std. dev.

Quartile measure is { (Q3 – Q2) – (Q2 – Q1) } / { Q3 – Q1 }

Complications

For a unimodal distribution (a distribution with a single peak), negative skew commonly indicates that the tail is on the left side of the distribution, and positive skew indicates that the tail is on the right. In cases where one tail is long but the other tail is fat, skewness does not obey a simple rule. For example, a zero value in skewness means that the tails on both sides of the mean balance out overall; this is the case for a symmetric distribution but can also be true for an asymmetric distribution where one tail is long and thin, and the other is short but fat. Thus, the judgement on the symmetry of a given distribution by using only its skewness is risky; the distribution shape must be taken into account.