Statistical Methods and Data Analysis I
Lecture 13: Distribution Moments, Skewness and Kurtosis.

Oleg Goldshmidt
oleg.goldshmidt@post.idc.ac.il

Arison School of Business
Interdisciplinary Center (IDC)
Herzliya, Israel

April 30, 2019

Central Moments

The n-th central moment of a real random variable

$$\mu_n = E\left[(x - E[x])^n\right] = \int_{-\infty}^{+\infty} (x-\mu)^n p(x) dx$$

- The *zeroth* central moment: $\mu_0 = 1$
- The *first* central moment: $\mu_1 = 0$ (*NB: not μ, which is called the first raw moment!*)
- The *second* central moment is the *variance*: $\mu_2 = \sigma^2$

Properties of central moments:

- Translation invariance: $\mu_n(x + c) = \mu_n(x)$
- Homogeneity of degree n: $\mu_n(cx) = c^n \mu_n(x)$
- Additivity, for independent random variables x and y, and *only* for $n \in \{1, 2, 3\}$: $\mu_n(x + y) = \mu_n(x) + \mu_n(y)$

Standardized Moments

Definition

The *standardized moment* of degree n of a real-valued random variable x with probability distribution $p(x)$ is the ratio of its n-th *central moment* μ_n to the n-th power of its *standard deviation* σ:

$$\tilde{\mu}_n = \frac{\mu_n}{\sigma^n} = \frac{E\left[(x-\mu)^n\right]}{\left(\sqrt{E\left[(x-\mu)^2\right]}\right)^n} = \int_{-\infty}^{+\infty} \left(\frac{x-\mu}{\sigma}\right)^n p(x)dx$$

Motivation: since *central moments are homogeneous of degree n*, $\mu_n(cx) = c^n \mu(x)$, the *standardized moments are scale-invariant*: $\tilde{\mu}_n(cx) = \tilde{\mu}_n(x)$. They are also *dimensionless*.

Special cases:

- The *first* standardized moment $\tilde{\mu}_1 = \mu_1/\sigma$ is always zero (as $\mu_1 \equiv 0$).
- The *second* standardized moment $\tilde{\mu}_2 = \mu_2/\sigma^2$ is always 1 (as $\mu_2 \equiv \sigma^2$).

Skewness: A Measure of Asymmetry

Definition: moment coefficient of skewness

The *third standardized moment* of the probability distribution of a real-valued random variable is called *skewness*:

$$\gamma_1 = \tilde{\mu}_3 = E\left[\frac{(x-\mu)^3}{\sigma^3}\right] = \frac{\mu_3}{\sigma^3}$$

Definition: mode skewness

$$\gamma_1 = \frac{mean - mode}{\sigma}$$

Definition: median skewness

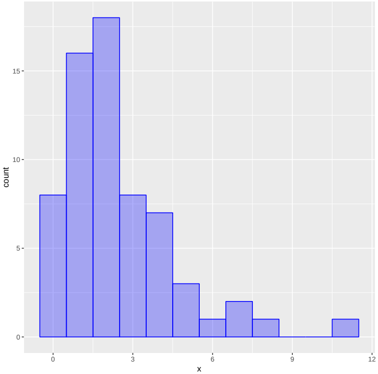
$$\gamma_1 = 3\frac{mean - median}{\sigma}$$

Definition: quartile-based skewness

$$\gamma_1 = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1} = \frac{\frac{Q_3+Q_1}{2} - Q_2}{\frac{Q_3-Q_1}{2}} = \frac{\text{"location"} - median}{\text{"dispersion"}}$$

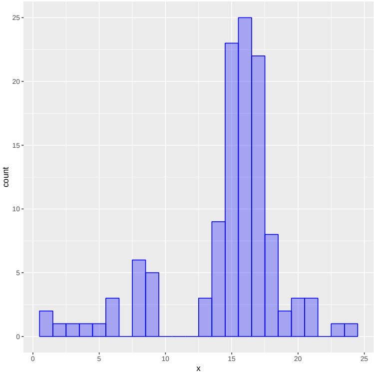
Skewness Examples

right skew



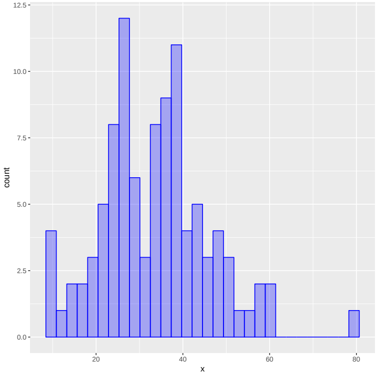
$\gamma_1 \approx 1.84$

left skew



$\gamma_1 \approx -1.35$

symmetric



$\gamma_1 \approx 0.51$

($\gamma_1 \approx 0.16$ without the outlier)

Kurtosis: A Measure of “Tailedness”

Definition: moment coefficient of kurtosis

The *fourth standardized moment* of the probability distribution of a real-valued random variable is called *kurtosis*:

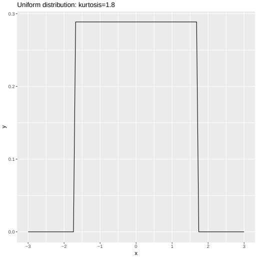
$$\kappa = \tilde{\mu}_4 = E\left[\left(\frac{x-\mu}{\sigma}\right)^4\right] = E\left[\frac{(x-\mu)^4}{\sigma^4}\right] = \frac{\mu_4}{\sigma^4}$$

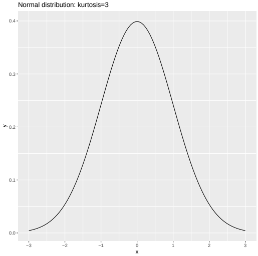
Interpretation: kurtosis measures how “heavy” or “fat” the distribution tails are. It says nothing of value about the peak: since it is the average of the standardized deviation to the 4th power, the values within 1 standard deviation from the mean contribute very little. Outliers contribute a lot.

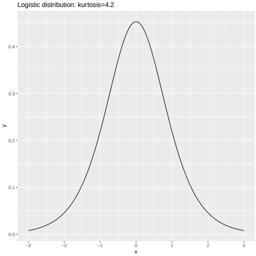
Note: The kurtosis of a normal distribution is 3 (calculate yourselves during your future mathematical analysis class). For this reason the definition of kurtosis is sometimes taken to be

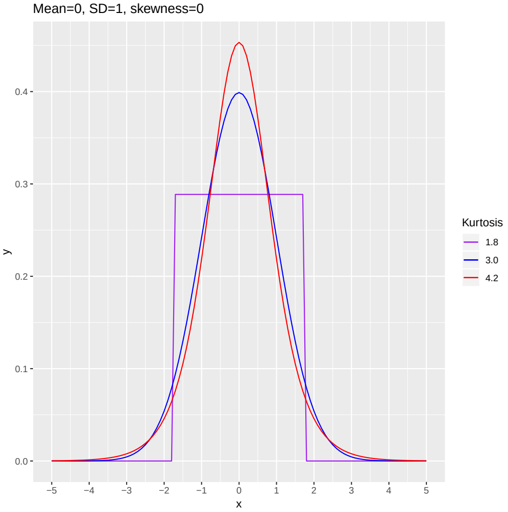
$$\kappa = \tilde{\mu}_4 - 3 = E\left[\frac{(x-\mu)^4}{\sigma^4}\right] - 3.$$

Kurtosis: Examples









Toolkit: Distribution Moments in R

```
> sample <- rnorm(10000,3,5)
> summary(sample)
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
-15.5309  -0.3519   3.0371   3.0345   6.5124  20.1189
> median(sample)
[1] 3.037139
> quantile(sample,seq(0.1,0.9,0.1))
       10%        20%        30%        40%        50%        60%        70%        80%        90%
-3.4381587 -1.1945945  0.3720708  1.7630074  3.0371387  4.3638092  5.7693382  7.3109338  9.4155437
> IQR(sample)
[1] 6.864276
> mean(sample)
[1] 3.034464
> var(sample)
[1] 25.36257
> sd(sample)
[1] 5.036126
> library(moments)
> skewness(sample)
[1] -0.01744459
> kurtosis(sample)
[1] 2.95811
```