Statistics for Data Science

Lecture 4

Moments

Ken Cameron

Content

- Moments
- Moment generating functions
- Leibnitz’s Rule

Moments

- If you are not a mathematician then moments are about any physical quantity that can be multiplied by distance.

- For example,

  - Force,

  - Electric charge,

  - Mass.

Mass

- 1st Moment - mass.
- 2nd Moment - centre of gravity.

Moments – Some Maths

- The n^{th} moment of a function f(x) about a value c

$$u_n = \int_{-\infty}^{\infty} (x-c)^n f(x)dx$$

c is usually zero.

Statistical Moments

- We can calculate moments for a set of data points.

$$\frac{(x_1^s + x_2^s + x_3^s + x_4^s + \ldots + x_n^s)}{n}$$

- Were s is the s^{th} moment.

First

- If s is 1:

$$\frac{(x_1 + x_2 + x_3 + x_4 + \ldots + x_n)}{n}$$

- 1st moment is the mean.

Second

- If s is 2:

$$\frac{(x_1^2 + x_2^2 + x_3^2 + x_4^2 + \ldots + x_n^2)}{n}$$

- Not that useful.

Second, about the mean

- Remember c, what if we make that the mean instead of zero

$$\frac{((x_1 - m)^2 + (x_2 - m)^2 + (x_3 - m)^2 + \ldots + (x_n - m)^2)}{n}$$

- 2nd moment about the mean is the variance. σ^2

Higher Moments

- Using the mean turns the moments into *central moments*.
  - We can also divide by the variance to normalise them.
  - So n^{th} normalised central moment is:

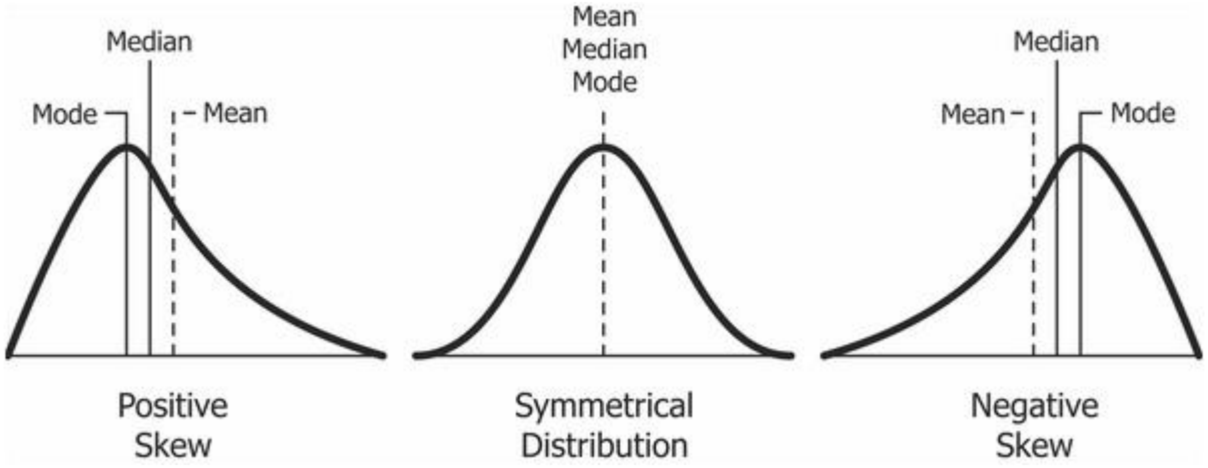
$$\frac{E[(X-\mu)^n]}{\sigma^2}$$

Third Moment

- Skewness

- A measure of the asymmetry of the probability distribution of a random variable (real valued) about the mean.

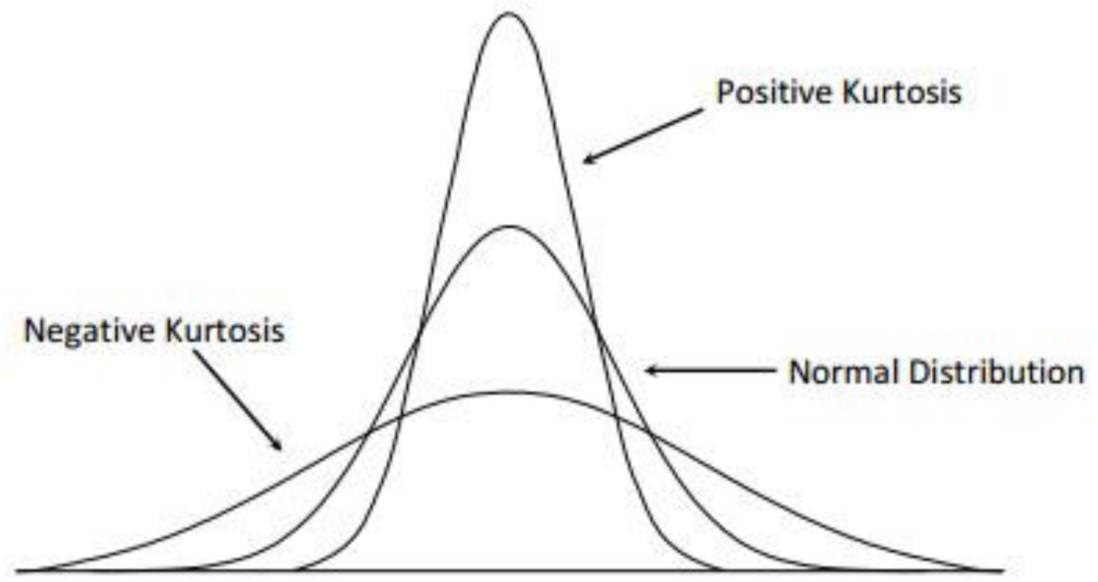
- Not always easy to interpret.



Fourth Moment

- Kurtosis

- Often described as a measure of ‘peakedness’ but is actually indicative of the tails.



Excess kurtosis

- Excess kurtosis is defines as kurtosis -3
- Mesokurtic. Distributions with zero excess kurtosis.
- Leptokurtic Positive excess kurtosis. (Fatter tails)
- Platykurtic Negative excess kurtosis. (Thinner tails.)

Some odd cases

- The 0^{th} moment about zero = 1.

- The 1^{st} moment about the mean = 0.

- Suppose X has an exponential distribution (λ),

- i.e. It has the pdf:

$$f_X(x) = \frac{1}{\lambda} e^{-x/\lambda}$$

0 <= x <= ∞, λ >= 0

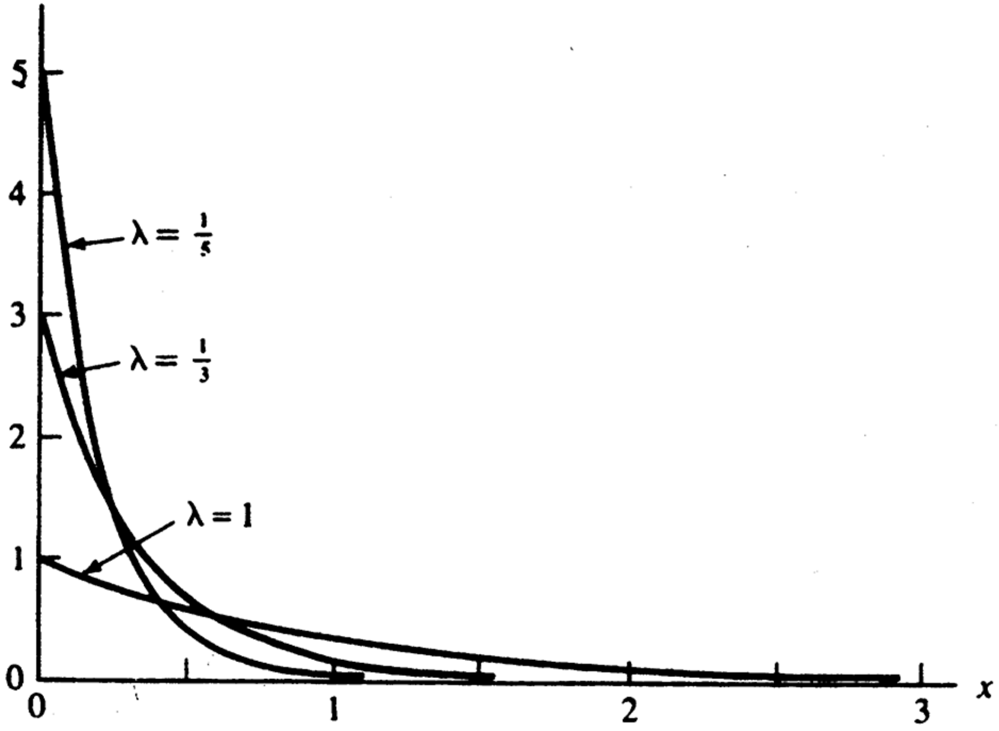
Calculate variance

- Var X = E(X - λ)^2 $= \int_0^\infty (x-\lambda)^2 \frac{1}{\lambda} e^{-x/\lambda} dx$

$$= \int_0^\infty (x^2 - 2x\lambda + \lambda^2) \frac{1}{\lambda} e^{-x/\lambda} dx$$

- Integrate by parts on the x^2 and x terms.

- Var X = λ^2


5
4
3
2
1
0
λ = 1/5
λ = 1/3
λ = 1
0
1
2
3
x

Moment generating functions

- Given a random variable X with a cdf F_x,

- We define a moment generating function of X as $M_x(t)$

$$M_x(t) = Ee^{tX}$$

- Provided the expectation exists for t in some neighbourhood of h.
  - I.e. $h > 0$ such that for all t, $-h < t < h$, Ee^{tX} exists.

mgf of X

$$M_x(t) = \sum_x e^{tx} P(X = x)$$

$$M_x(t) = \int_{-\infty}^{\infty} e^{tx} f_x(x) dx$$

Discrete

Continuous

If X has a mgf $M_x(t)$

then $EX^n = M_x^{(n)}(0)$

$$M_x^{(n)}(t) = \frac{d^n}{dt^n} Mx(t)\Big|_{t=0}$$

Proof

Assuming we can differentiate under the integral sign,

$$\frac{d}{dt} Mx(t) = \frac{d}{dt} \int_{-\infty}^{\infty} e^{tx} fx(x) dx$$

$$= \int_{-\infty}^{\infty} (\frac{d}{dt} e^{tx}) fx(x) dx$$

$$= \int_{-\infty}^{\infty} (xe^{tx}) fx(x) dx$$

$= EXe^{tx}$

Proof cont.

Thus,

$$\frac{d}{dt} Mx(t)\Big|_{t=0} = EXe^{tx}\Big|_{t=0} = EX$$

Analogously,

$$\frac{d^n}{dt^n} Mx(t)\Big|_{t=0} = EX^n e^{tx}\Big|_{t=0} = EX^n$$

Leibnitz’s Rule

- The proof required we interchange the order of the integration and differentiation.

- The more you look at theoretical statistics, the more often you will come across this requirement.

- So this raises the question, can we do this?

  - And under what conditions.

Leibnitz’s Rule

- The answer is yes.
  - Obviously, as we’ve already seen one example.
- The question of the conditions remains.
  - They can be established using standard theorems from calculus.
  - The detailed proofs will not be presented here.
  - You’ll find them in most calculus texts.

Leibnitz’s Rule

- Here, we’ll establish the method, Leibnitz’s Rule, of calculating:

$$\frac{d}{dx}\int_{a(x)}^{b(x)} f(x,t)\,dt$$

- Where $-\infty < a(x), b(x) < \infty$ for all x.

Leibnitz’s Rule

- If f(x,t), a(x) and b(x) are differentiable with respect to x then,

$$\frac{d}{dx}\left(\int_{a(x)}^{b(x)} f(x,t)\,dt\right) = f\big(x,b(x)\big)\cdot\frac{d}{dx}b(x) - f\big(x,a(x)\big)\cdot\frac{d}{dx}a(x) + \int_{a(x)}^{b(x)} \frac{\partial}{\partial x} f(x,t)\,dt,$$

- If a(x) and b(x) are constants:

$$\frac{d}{dx}\left(\int_{a}^{b} f(x,t)\,dt\right) = \int_{a}^{b} \frac{\partial}{\partial x} f(x,t)\,dt.$$

Limits

- It's the interchange of the limits that controls when Leibniz's Rule applies.

  - In general, given two limiting operations we cannot assume the same result is produced when the are applied in either order.

  - However,

  - If both upper and lower limits are taken as constants
  - and D_x is the partial derivative with respect to x,
  - and I_t is the integral operator with respect to t over a fixed interval

$$I_t D_x = D_x I_t$$