MOMENTS

Defn: Let X be a r.v. For each integer k, the k^{th} moment about the origin of X is

$$\mu'_k = \mathbb{E}X^k$$

and the k^{th} moment about the mean (a.k.a. "central moment") of X is

$$\mu_k = \mathbb{E}[X-\mu]^k,$$

where $\mu := \mu'_1$.

Remark: $\operatorname{Var} X = \mu_2 =: \sigma^2$ is the 2nd moment about the mean

* We borrow the term "moment" from physics, where in classical mechanics, a moment is the turning effect of a force.

* Other quantities involving moments:

Skewness: $\operatorname{Skew} X = \mathbb{E}\left[\frac{X-\mu}{\sigma}\right]^3 = \frac{\mu_3}{\sigma^3}$

Kurtosis: $\operatorname{Kurt} X = \mathbb{E}\left[\frac{X-\mu}{\sigma}\right]^4 = \frac{\mu_4}{\sigma^4}$

(Measure of "heavy-tailed-ness")

MOMENT-GENERATING FUNCTIONS

Defn: The moment-generating function (mgf) M_X of a r.v. X is

$$M_X(t) = \mathbb{E} \exp[tX],$$

provided the expectation is finite for t in a neighborhood of 0, that is, provided there exists a $C > 0$ such that $M_X(t) < \infty$ for all $-C < t < C$.

In particular,

$$M_X(t) = \begin{cases} \int_{-\infty}^{\infty} e^{tx} f_X(x)\, dx & \text{if } X \text{ cont w/ pdf } f_X \\ \sum_{x \in \mathcal{X}} e^{tx} p_X(x) & \text{if } X \text{ disc. w/ pmf } p_X. \end{cases}$$

"A moment-generating function is a function that generates moments."
— J. Tebbs

* How does the mgf generate moments?

Theorem: If X is a r.v. with mgf M_X, then

$$\mathbb{E}X^k = M_X^{(k)}(0),$$

where

$$M_X^{(k)}(0) = \left(\frac{d}{dt}\right)^k M_X(t) \Big|_{t=0}.$$

This means we evaluate the foregoing expression at $t=0$

* Thus, to get $\mu_k = \mathbb{E}X^k$, we find the k^{th} derivative of $M_X(t)$ and evaluate it at $t=0$.

Proof: Begin with a Taylor expansion of e^{tx} around $x=0$, writing

$$e^{tx} = \sum_{j=0}^{\infty} \frac{t^j}{j!} x^j.$$

$$\boxed{g(x) = \sum_{j=0}^{\infty} \frac{g^{(j)}(x_0)(x-x_0)^j}{j!}}$$

Provided $\mathbb{E}X^j < \infty$ for all $j=1,2,\ldots$, we have

$$M_X(t) = \sum_{j=0}^{\infty} \frac{t^j}{j!} \mathbb{E}X^j$$

$$= 1 + t\,\mathbb{E}X + \frac{t^2}{2}\mathbb{E}X^2 + \frac{t^3}{3!}\mathbb{E}X^3 + \cdots$$

We see that the k^{th} derivative with respect to t is

$$M_X^{(k)}(0) = \mathbb{E}X^k + t\,\mathbb{E}X^{k+1} + \frac{t^2}{2}\mathbb{E}X^{k+2} + \frac{t^3}{3!}\mathbb{E}X^{k+3} + \cdots$$

So $M_X^{(k)}(0) = \mathbb{E}X^k$.

E.g. Mgf of $X \sim \text{Exponential}(\lambda)$.

We have

$$\mathbb{E}\, e^{tX} = \int_0^\infty e^{tX} \cdot \frac{1}{\lambda} e^{-x/\lambda}\, dx$$

$$= \int_0^\infty \frac{1}{\lambda} e^{-x\left(\frac{1}{\lambda} - t\right)}\, dx$$

$$= \frac{1}{\lambda}\left(\frac{\lambda}{1-\lambda t}\right) \underbrace{\int_0^\infty \left(\frac{\lambda}{1-\lambda t}\right)^{-1} e^{-x/\left(\frac{\lambda}{1-\lambda t}\right)}\, dt}_{\text{Integrates to 1 as Exponential}\left(\frac{\lambda}{1-\lambda t}\right)\text{ pdf, provided } t < 1/\lambda}$$

Note: $\frac{\lambda}{1-\lambda t} > 0 \iff 1-\lambda t > 0 \iff t < 1/\lambda$

$$= (1-\lambda t)^{-1}.$$

So $\quad M_X(t) = (1-\lambda t)^{-1} \quad \text{for} \quad t < \frac{1}{\lambda}$

<u>Use it</u>:

$$\mathbb{E}X = \frac{d}{dt}(1-\lambda t)^{-1}\Big|_{t=0} = -(1-\lambda t)^{-2}(-\lambda)\Big|_{t=0} = \lambda.$$

$$\mathbb{E}X^2 = \left(\frac{d}{dt}\right)^2 (1-\lambda t)^{-1}\Big|_{t=0} = -2\lambda(1-\lambda t)^{-3}(-\lambda)\Big|_{t=0} = 2\lambda^2.$$

So

$$\text{Var } X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = 2\lambda^2 - \lambda^2 = \lambda^2.$$

E.g. Mgf of $X \sim \text{Gamma}(\alpha, \beta)$

$$M_X(t) = \int_0^\infty \exp[tx] \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp\left[-\frac{x}{\beta}\right] dx$$

$$= \int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp\Big[\underbrace{-\frac{x}{\beta} + tx}_{= -x\left(\frac{1}{\beta} - t\right) = -x\left(\frac{1-\beta t}{\beta}\right) = \frac{-x}{\left(\frac{\beta}{1-\beta t}\right)}}\Big] dx$$

$$= \left(\frac{\beta}{1-\beta t}\right)^\alpha \frac{1}{\beta^\alpha} \underbrace{\int_0^\infty \frac{1}{\Gamma(\alpha)\left(\frac{\beta}{1-\beta t}\right)^\alpha} x^{\alpha-1} \exp\left[-x\Big/\left(\frac{\beta}{1-\beta t}\right)\right] dx}_{= \begin{cases} 1 & \text{if } t < \frac{1}{\beta} \\ \infty & \text{if } t \geq \frac{1}{\beta} \end{cases}}$$

$$= (1-\beta t)^{-\alpha} \quad \text{for} \quad t < \frac{1}{\beta}$$

Use it:

$$\mathbb{E}X = \frac{d}{dt}(1-\beta t)^{-\alpha}\Big|_{t=0}$$

$$= (-\alpha)(1-\beta t)^{-\alpha-1}(-\beta)\Big|_{t=0}$$

$$= \alpha\beta$$

$$\mathbb{E}X^2 = \left(\frac{d}{dt}\right)^2(1-\beta t)^{-\alpha}\Big|_{t=0}$$

$$= \frac{d}{dt}\,\alpha\beta\,(1-\beta t)^{-\alpha-1}\Big|_{t=0}$$

$$= \alpha\beta(-\alpha-1)(1-\beta t)^{-\alpha-2}(-\beta)\Big|_{t=0}$$

$$= \alpha\beta^2(\alpha+1)$$

So $\quad \mathrm{Var}\, X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \alpha\beta^2(\alpha+1) - (\alpha\beta)^2 = \alpha\beta^2.$

E.g. Mgf of $X \sim \text{Binomial}(n,p)$

$$M_X(t) = \sum_{x=0}^{n} e^{tx}\binom{n}{x}p^x(1-p)^{n-x}$$

$$= \sum_{x=0}^{n}\binom{n}{x}(pe^t)^x(1-p)^{n-x}$$

$$= \left[pe^t + (1-p)\right]^n$$

$\boxed{\text{Binomial Theorem: for any real numbers } a \text{ and } b,\quad (a+b)^n = \sum_{x=0}^{n}\binom{n}{x}a^x b^{n-x}}$

Use it:

$$\mathbb{E}X = \frac{d}{dt}\left[pe^t + (1-p)\right]^n \Big|_{t=0}$$

$$= n\left[pe^t + (1-p)\right]^{n-1} pe^t \Big|_{t=0}$$

$$= np$$

$$\mathbb{E}X^2 = \left(\frac{d}{dt}\right)^2 \left[pe^t + (1-p)\right]^n \Big|_{t=0}$$

$$= \frac{d}{dt}\, n\left[pe^t + (1-p)\right]^{n-1} pe^t \Big|_{t=0}$$

$$= n(n-1)\left[pe^t + (1-p)\right]^{n-2}\left(pe^t\right)^2 + n\left[pe^t + (1-p)\right]^{n-1} pe^t \Big|_{t=0}$$

$$= n(n-1)p^2 + np$$

$$= (np)^2 + np(1-p)$$

So $\text{Var } X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = (np)^2 + np(1-p) - (np)^2 = np(1-p)$

ACTUALLY: Mgfs are more useful for characterizing distributions than for computing moments.

Theorem: Let $X \sim F_X$ and $Y \sim F_Y$ and suppose $\mathbb{E}X^k < \infty$ and $\mathbb{E}Y^k < \infty$ for all $k = 1, 2, \ldots$

Then if M_X and M_Y exist and $M_X(t) = M_Y(t)$ for all t in a neighborhood of 0, then $F_X = F_Y$; that is X and Y are identically distributed.

* R.v.s with the same mgf have same dist.

* We will later use mgfs to prove a version of the Central Limit Theorem. Basically:

  If $\bar{X}$ is the mean of a sample drawn from a distribution with an mgf, then

  $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ behaves more and more like $Z \sim \text{Normal}(0,1)$

  as $n \to \infty$. We can show this by showing that

  $$\lim_{n\to\infty} M_{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}(t) = M_Z(t)$$

  But we aren't quite ready for this yet.

  But let's find M_Z:

  $$M_Z(t) = \int_{-\infty}^{\infty} \exp[tz] \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right] dz$$

  $$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[\underbrace{-\frac{z^2}{2} + tz}\right] dz$$

"complete the square": $-\frac{z^2}{2} + tz = -\frac{z^2 - 2tz}{2} = -\frac{z^2 - 2tz + t^2 - t^2}{2} = -\frac{(z-t)^2}{2} + \frac{t^2}{2}$

  $$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(z-t)^2}{2} + \frac{t^2}{2}\right] dz$$

  $$= \exp\left[t^2/2\right] \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(z-t)^2}{2}\right] dz}_{=1\text{, integral over pdf of Normal}(t,1)}$$

  $$= e^{t^2/2}$$

* The following theorem will let us write M_X for $X \sim \text{Normal}(\mu, \sigma^2)$ in terms of M_Z.

Theorem: For any constants a and b, the mgf of $aX+b$ is

$$M_{aX+b}(t) = e^{tb} M_X(at)$$

Proof:

$$M_{aX+b}(t) = \mathbb{E} \exp\left[t(aX+b)\right] = \mathbb{E} \exp\left[(ta)X\right] \exp\left[tb\right] = e^{tb} M_X(at)$$

E.g. Find the mgf of $X \sim \text{Normal}(\mu, \sigma^2)$.

We can write $X = \mu + \sigma Z$, where $Z \sim \text{Normal}(0,1)$.

Then $M_X(t) = M_{\mu+\sigma Z}(t) = e^{\mu t} M_Z(\sigma t)$,

So $M_X(t) = e^{t\mu} e^{(\sigma t)^2/2} = \exp\left[t\mu + \frac{\sigma^2 t^2}{2}\right]$

Use it:

$$\mathbb{E}X = \frac{d}{dt} \exp\left[t\mu + \frac{\sigma^2 t^2}{2}\right]\Bigg|_{t=0}$$

$$= \exp\left[t\mu + \frac{\sigma^2 t^2}{2}\right](\mu + \sigma^2 t)\Bigg|_{t=0}$$

$$= \mu$$

$$\mathbb{E}X^2 = \left(\frac{d}{dt}\right)^2 \exp\left[t\mu + \frac{\sigma^2 t^2}{2}\right]\Bigg|_{t=0}$$

$$= \frac{d}{dt} \exp\left[t\mu + \frac{\sigma^2 t^2}{2}\right](\mu + \sigma^2 t)\Bigg|_{t=0}$$

$$= \exp\left[t\mu + \frac{\sigma^2 t^2}{2}\right](\mu + \sigma^2 t)^2 + \exp\left[t\mu + \frac{\sigma^2 t^2}{2}\right]\sigma^2 \Bigg|_{t=0}$$

$$= \mu^2 + \sigma^2,$$

So $\text{Var}\,X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \mu^2 + \sigma^2 - \mu^2 = \sigma^2.$