2.3 Moments

Def. For each $n \in \mathbb{N}$, the n-th moment of X, or $F_X(x)$, is

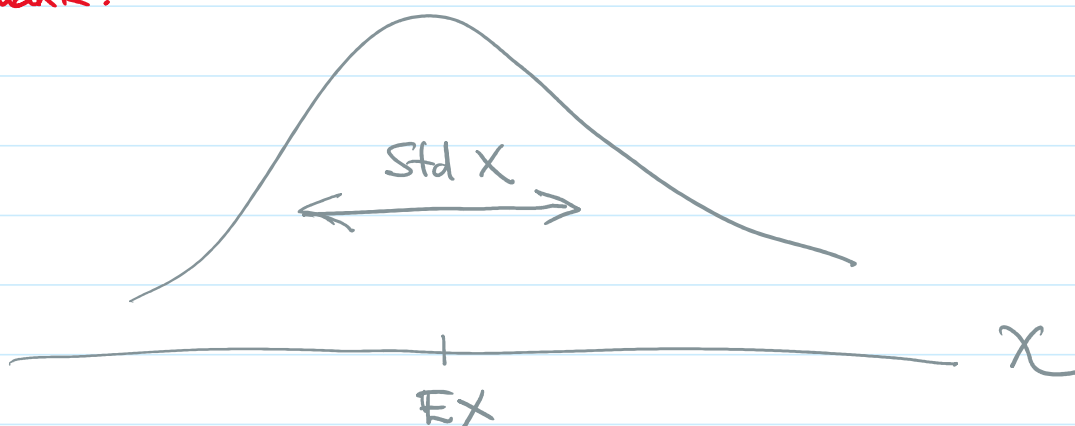$$\mu'_n := \mathbb{E}X^n$$

The n-th central moment of X is

$$\mu_n := \mathbb{E}(X - \mathbb{E}X)^n$$

Let $n = 2$: The second central moment is known as variance

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

Remark.

- $\mathbb{E}X$ captures the location of X's distribution
- $\text{Var}X$, or $\text{Std }X$ capture the spread.

Why?

least square

Consider the $\overset{\vee}{\text{optimization}}$ $\quad \min_{b \in \mathbb{R}} \mathbb{E}(X-b)^2.$

$$\mathbb{E}(X-b)^2$$
$$= \mathbb{E}\, X^2 - 2bX + b^2$$
$$= (\mathbb{E}\, X^2) - 2b\,\mathbb{E}X + b^2 \qquad \text{(linearity of expect.)} \quad \text{L.O.E.}$$

$$= b^2 - 2b(\mathbb{E}X) + (\mathbb{E}X)^2 + \underbrace{\mathbb{E}X^2 - (\mathbb{E}X)^2}_{\text{Var}X}$$
$$= (b - \mathbb{E}X)^2 + \qquad\qquad\qquad\qquad \leftarrow \text{see } (*)$$

$$\geqslant \text{Var}\, X.$$

The "location b" which solves the least square problem is

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

$(*)$
$$\text{Var}\, X = \mathbb{E}(X - \mathbb{E}X)^2$$
$$= \mathbb{E}\left(X^2 - 2X \cdot \mathbb{E}X + (\mathbb{E}X)^2\right)$$
L.O.E. $\rightsquigarrow$
$$= \mathbb{E}X^2 - 2\,\mathbb{E}X \cdot \mathbb{E}X + (\mathbb{E}X)^2$$
$$= \mathbb{E}X^2 - (\mathbb{E}X)^2$$

Example - Exponential dist.

Recall: $X \sim Exp(\lambda)$ has $\mathbb{E}X = \lambda$

$$var\, X = \mathbb{E}X^2 - (\mathbb{E}X)^2$$

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

integration by part

$$\stackrel{=}{=} 2\lambda^2 - \lambda^2$$

$$= \lambda^2$$

Recall $\quad \mathbb{E}(aX+b) = a\,\mathbb{E}X + b$

Thm. if X is a Random variable with finite variance. Then for any $a, b \in \mathbb{R}$

$$\boxed{\text{Var}(aX+b) = a^2 \text{Var}\,X.}$$

Proof.

$$\text{var}(aX+b) = \mathbb{E}\left[(aX+b)^2 - (a\,\mathbb{E}X + b)^2\right]$$

$$= \mathbb{E}\left[a^2X^2 + 2aXb + b^2 - a^2(\mathbb{E}X)^2 - 2ab\,\mathbb{E}X - b^2\right]$$

L.D.E. $\quad = a^2\,\mathbb{E}X^2 + 2ab\,\mathbb{E}X - a^2(\mathbb{E}X)^2 - 2ab\,\mathbb{E}X$

$$= a^2\left(\mathbb{E}X^2 - (\mathbb{E}X)^2\right)$$

before $\quad = a^2\,\text{Var}\,X.$ $\qquad\qquad\qquad \square$

Example - Binomial dist.

Recall   if $X \sim$ Binomial (n, p) $n \in \mathbb{N}, p \in (0,1)$

then $\mathbb{E}X = np$

$$\mathbb{E}X^2 = \sum_{x=0}^{n} x^2 f_X(x)$$

$$= \sum_{x=0}^{n} x^2 \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=1}^{n} xn \binom{n-1}{x-1} p^x (1-p)^{n-x}$$

$y = x-1$

$$= \sum_{y=0}^{n-1} n(y+1) \binom{n-1}{y} p^{y+1} (1-p)^{n-y-1}$$

$$= np \cdot \left[\sum_{y=0}^{n-1} y \binom{n-1}{y} p^y (1-p)^{n-1-y} + \sum_{y=0}^{n-1} \binom{n-1}{y} p^{y} (1-p)^{n-1-y} \right]$$

$$= np \left[(n-1) p + 1 \right]$$

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

Hence $\text{var } X = \mathbb{E}X^2 - (\mathbb{E}X)^2$

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

$$= np(1-p)$$

FAST WAY:

Let $X_i = \begin{cases} 1 & \text{if the } i\text{-th toss is a head} \\ 0 & \text{otherwise} \end{cases}$

Then $\mathbb{E}X_i = p \cdot 1 + (1-p) \cdot 0 = p.$

$\mathbb{E}X_i^2 = \mathbb{E}X_i = p.$

$\Rightarrow \text{var } X_i = \mathbb{E}X_i^2 - (\mathbb{E}X)^2$

$\qquad\qquad\qquad = p - p^2 = p(1-p)$

Since $X = X_1 + \ldots + X_n$

and $X_1, \ldots, X_n$ are mutually independent

then we'll learn later that

$\text{var } X = \text{var } X_1 + \ldots + \text{var } X_n$

$\qquad\quad = n \times \text{var } X_1$

$\qquad\quad = n\, p(1-p).$

MOMENT GENERATING FUNCTIONS (MGF)

Def. Let X be a R.V. with cdf F_X

The moment generating function (mgf) is

$$M_X(t) := \mathbb{E}\, e^{tX}$$

provided that the expectation exists in some neighborhood of $t = 0$ (i.e. for $t \in$ some $(-h, h)$)
$h > 0$

Remark

○ $M_X(t) = \begin{cases} \int e^{tx} f_X(x)\, dx & \text{for cont. } X \\ \sum_x e^{tx} f_X(x) & \text{for disc. } X \end{cases}$

○ The mgf is can be used to "characterize" the distribution (of the Random variable)

Thm. $\forall n \in \mathbb{N}$ $\boxed{\mathbb{E}X^n = \left.\dfrac{d^n}{dt^n} M_X(t)\right|_{t=0}}$

Remark.

o we also write RHS as $M_X^{(n)}(0)$
(n-th derivative of the mgf evaluated
at $t=0$ gives the n-th moment).

o Thm gives the meaning of " moment generating "

Proof.

$n=0$ $\mathbb{E}X^0 = \mathbb{E}1 = 1$

$M_X(0) = \mathbb{E}e^{0 \cdot X} = \mathbb{E}1 = 1$.

$\leftarrow$ or Σ

$n=1$ $\dfrac{d}{dt} M_X(t) = \dfrac{d}{dt} \int e^{tx} f_X(x) \, dx$

$= \int \dfrac{d}{dt} e^{tx} f_X(x) \, dx$ (exchanging integral and differentiation)

$= \int x \, e^{tx} f_X(x) \, dx$

$\Rightarrow \left.\dfrac{d}{dt} M_X(t)\right|_{t=0} = \int x f_X(x) \, dx = \mathbb{E}X.$

$$n=2 \qquad \frac{d^2}{dt^2} M_X(t) = \int \frac{d}{dt} \, x e^{tx} f_X(x) \, dx$$

$$= \int x^2 e^{tx} f_X(x) \, dx$$

$$\Rightarrow \quad \frac{d^2}{dt^2} M_X(t) \Big|_{t=0} = \mathbb{E}X^2.$$

And so on, for any $n \in \mathbb{N}$. $\qquad \square$

Example   (Gamma)

For $\alpha, \beta > 0$ define Gamma function

$$\Gamma(\alpha) = \frac{1}{\beta^\alpha} \int_0^\infty x^{\alpha-1} e^{-x/\beta} \, dx \qquad (*)$$

NOTE : $\Gamma(\alpha)$ does not depend on β, because by change of variable

$y = \frac{x}{\beta}$, $\quad \frac{1}{\beta^\alpha} \int_0^\infty x^{\alpha-1} e^{-x/\beta} \, dx = \frac{1}{\beta^\alpha} \int_0^\infty (\beta y)^{\alpha-1} e^{-y} \beta \, dy$

$$= \int_0^\infty y^{\alpha-1} e^{-y} \, dy$$

The last expression doesn't depend on β, and usually is taken as the <u>definition</u> of $\Gamma(\alpha)$.

Define $\quad f(x) := \frac{1}{\Gamma(\alpha) \beta^\alpha} x^{\alpha-1} e^{-x/\beta}$, $\quad x \in (0, +\infty)$

f is a pdf on $(0, +\infty)$. This is indeed the pdf ~~for~~ Gamma distribution, denoted Gamma (α, β)

For $X \sim \text{Gamma}(\alpha, \beta)$

$$M_X(t) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty e^{tx} x^{\alpha-1} e^{-x/\beta} \, dx$$

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

$$\searrow \frac{1-\beta t}{\beta}$$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \cdot \Gamma(\alpha) \left(\frac{\beta}{1-\beta t}\right)^\alpha \quad , \text{ if } 1-\beta t > 0$$

$$= \left(\frac{1}{1-\beta t}\right)^\alpha .$$

if $t \geqslant \frac{1}{\beta}$ then $\frac{1}{\beta} - t \leqslant 0$, then $\int_0^\infty x^{\alpha-1} e^{-x\left(\frac{1}{\beta} - t\right)} \, dx = \infty$

$$M_X(t) = +\infty .$$

Now,

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

$$= \alpha\beta .$$

Example (Binomial)

Let $X \sim$ Binomial (n, p)

$$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad , x = 0, 1, \ldots, n$$

Then

$$
\begin{aligned}
M_X(t) &= \mathbb{E}\, e^{tX} \\
&= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x} \\
&= \sum_{x=0}^{n} \binom{n}{x} (e^t p)^x (1-p)^{n-x} \\
&= (e^t p + 1-p)^n \qquad \leftarrow \text{ binomial formula}
\end{aligned}
$$

Differentiate $M_X(t)$ upto k-th order and evaluate it at $t=0$ to obtain $\mathbb{E} X^k$!

Remark

- The mgf uniquely determines all moments
- does the (infinite) set of moments uniquely determine / identify the distribution ?