Moment Generating Function (MGF) :-

X : Random variable

$M_X(t) = E[e^{tX}]$ is called the MGF if the expectation exists in some neighbourhood of the origin.

MGF of the distribution associated with the r.v. X

Remarks

1. If MGF exists, it then characterises the distribution.

Mathematically, $\underset{\text{r.v.}}{X} \sim \underset{\text{CDF}}{F}$ & $\underset{\text{r.v.}}{Y} \sim \underset{\text{CDF}}{G}$

Result $\quad M_X(t) = M_Y(t) \;\; \forall t$

$\iff X \stackrel{d}{=} Y.$ (X & Y will be identically distributed)

2. $M_X(t)$ generates the moments.

Observe that

Not trivial

$$\frac{d^k}{dt^k} M_X(t)\Big|_{t=0} = \frac{d^k}{dt^k} E(e^{tX})\Big|_{t=0} = E\left[\frac{d^k}{dt^k} e^{tX}\right]\Big|_{t=0}$$

$$= E\left[X^k e^{tX}\Big|_{t=0}\right] = E(X^k)$$

That's why, $M_X(t)$ is called the Moment generating f^n.

3. If $M_X(t)$ exists, one can think according to the following way.

$$M_X(t) = M_X(0) + \frac{t}{1!} \times \frac{d}{dt} M_X(0) + \frac{t^2}{2!} M_X''(0) + \frac{t^3}{3!} M_X'''(0) + \ldots$$

Observe that the coeff. of $\frac{t^k}{k!}$ is $M_X^{\overbrace{'''''}^{k \text{ times}}}(t)\Big|_{t=0} = E(X^k)$.

4) MGF may NOT exist always.

Ex:- $X \sim$ Cauchy $(0,1) \longrightarrow f_X(x) = \frac{1}{\pi(1+x^2)}$, $x \in \mathbb{R}$.

X ↓ r.v

Here MGF does NOT exist.

Infact, any moment (of order ≥ 1) do not exist. ↳ E.g. $E(X)$ or $E(X^6)$ does not exist.

However, the moments (of order < 1) exist (e.g. $E(X^{3/4})$ or $E(X^{\frac{7}{8}})$ exists).

5) Since MGF characterises the dist$^{\underline{n}}$s, it can be used to formulated test-statistic for 2 sample problem (ie, $X \sim F$ & $Y \sim G$, $H_0: F = G$ ag. $H_1: F \neq G$) or 2 sample supervised learning problem.

Two important inequalities:-

1. Chebyshev's inequality:-

X : Random variable

$E(X) := \mu$

$Var(X) = E[X - E(X)]^2 := \sigma^2$

Then for any $\varepsilon > 0$, we have

$$P[|X - \mu| \geq \varepsilon] \leq \frac{\sigma^2}{\varepsilon^2}.$$

Proof:- Given $\varepsilon > 0$,

$$\sigma^2 = E[X-\mu]^2 = \int_{-\infty}^{\infty} (x-\mu)^2 \underbrace{f_X(x)}_{\text{p.d.f. of } X} \, dx$$

$$= \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)\, dx$$

$$= \int_{\{x : |x-\mu| > \varepsilon\}} (x-\mu)^2 f_X(x)\, dx + \int_{\{x : |x-\mu| \leq \varepsilon\}} (x-\mu)^2 f_X(x)\, dx .$$

$$\geq \int_{|x-\mu| \geq \varepsilon} (x-\mu)^2 f_X(x)\, dx$$ since $(x-\mu)^2 \geq 0$ & $f_X(x) \geq 0$ & $\{x : |x-\mu| \leq \varepsilon\} \geq \phi$.

$$\geq \varepsilon^2 \int_{\{x : |x-\mu| \geq \varepsilon\}} f_X(x)\, dx$$

$$= \varepsilon^2 P[\omega : |X(\omega) - \mu| \geq \varepsilon] \Longleftrightarrow P[|X-\mu| \geq \varepsilon] \leq \frac{\sigma^2}{\varepsilon^2} .$$

<u>Comment</u>:- Such inequality helps us to convergence in prob for many cases.

Markov's inequality (generalization)

If $h(X)$ is a non-negative f^n & $E[h(X)] < \infty$,

then $P[h(X) > \varepsilon] \leq \dfrac{E[h(X)]}{\varepsilon}$ $\forall\ \varepsilon > 0$.

Proof:- Same proof will work [consider $h(X)$ instead of $(X-\mu)^2$].

Remark:- $1_A = 1$ if A is true

$= 0$, otherwise.

$P[A] = E 1_A$ $\left[E 1_A = 1 \times P[A] + 0 \times P(A^c) = P(A) \right]$.

A is an event

Rough work.

$\Big[P[h(X) > \varepsilon] \leq \dfrac{E h(X)}{\varepsilon}$, $\quad h(X) 1_{\{h(X) \geq \varepsilon\}} \geq \varepsilon \Leftrightarrow E h(X) \geq \varepsilon P[h(X) \geq \varepsilon]$

$\Leftrightarrow E[h(X) 1_{\{h(X) \geq \varepsilon\}}] \geq \varepsilon$. $\Big]$

Next concepts :-

Median & Quantiles :-

Median :- Middle most observation

Population median :-

X : Random variable

F : CDF

Median of X is defined as $F^{-1}(\frac{1}{2})$ if F is cont^{n}.

$F^{-1}(\frac{1}{2})$: If X is cont^{n}

$$F^{-1}(\frac{1}{2}) : \inf_{x} \left\{ x : F(x) \geq \frac{1}{2} \right\}.$$