Probability and statistics

(September-14)

Expectation of a r.v.

X is a discrete r.v. with p.m.f. $f(x)$

then $E(X) = \sum_{x \in R_X} x f(x)$

i) $X \sim$ Bernoulli(p) ; $E(X) = p$

ii) $X \sim$ Binomial (n, p) ; $E(X) = np$

iii) $E(X) = \sum_{x \in R_X} x f(x)$ exists only if $\sum_{x \in R_X} x f(x)$ is covergent.

Expectation of a fn of discrete r.v.

Let $\underbrace{z = \varphi(x)}$ be a fn of discrete r.v. X

$$E(Z) = E(\varphi(X)) = \sum_{x \in R_X} \varphi(x)\, f(x)$$

provided that the expectation exists.

Ex:

x	-2	-1	0	1	2
$f(x)$	$1/5$	$1/5$	$1/5$	$1/5$	$1/5$

x^2	4	1	0	1	4
$f(x)$	$1/5$	$1/5$	$1/5$	$1/5$	$1/5$

$z = \varphi(x) = x^2$

$E(X) = 0$

$$E(X^2) = \sum_{x \in R_X} x^2 f(x) = 4 \times \frac{1}{5} + 1 \times \frac{1}{5} + 0 \times \frac{1}{5} + 1 \times \frac{1}{5} + 4 \times \frac{1}{5}$$

$$= 2$$

Properties of expectation:

Let X & Y be discrete r.v.s with finite expectations.

i) For any $\alpha \in \mathbb{R}$ s.t. $P(X=\alpha)=1$; $E(X)=\alpha$

ii) for some constant $c \in \mathbb{R}$, $E(cX) = c\,E(X)$

iii) $E(X+Y) = E(X) + E(Y)$

iv) If $P(X \geq Y) = 1$, then $E(X) \geq E(Y)$

v) $|E(X)| \leq E(|X|)$

$$\left| \frac{x_1 + x_2 + \cdots + x_n}{n} \right|$$

$$\leq \frac{|x_1|}{n} + \frac{|x_2|}{n} + \cdots + \frac{|x_n|}{n}$$

x	-2	-1	0	1	2
$f(x)$	$1/5$	$1/5$	$1/5$	$1/5$	$1/5$

Moment generating function (MGF)

X is a discrete r.v. with pmf $f(x)$.

Consider this fn $z = \varphi(x) = e^{tx}$

$$\boxed{M_X(t) = E(e^{tX})} \quad \ldots \text{ MGF of r.v. } X.$$

$$= \sum_{\underbrace{x_i \in R_X}} e^{t x_i} f(x_i)$$

Ex: $X \sim \text{Bernoulli}(p)$

x	0	1
$f(x)$	$1-p$	p

$$M_X(t) = e^{t \cdot 0}(1-p) + e^{t \cdot 1} p$$

$$\boxed{M_X(t) = (1-p) + pe^{t}}$$

Ex: $X \sim \text{Binomial}(n, p)$

$$f(x) = \binom{n}{x} p^{x} (1-p)^{n-x} \qquad x = 0, 1, 2, \dots, n$$

$$= 0 \qquad \text{o.w.}$$

$$M_X(t) = E(e^{tX})$$

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

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

$$\boxed{M_X(t) = \left((1-p) + pe^{t}\right)^{n}}$$

Ex: $X \sim \text{Poission}(\lambda)$

$$f(x) = \frac{e^{-\lambda} \lambda^x}{x!} \qquad x = 0, 1, 2, \ldots$$

$$= 0 \qquad \text{o.w.}$$

$$M_X(t) = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda} \lambda^x}{x!}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(e^t \lambda)^x}{x!}$$

$$\boxed{M_X(t) = e^{-\lambda} e^{\lambda e^t}}$$

Moments:

For a discrete r.v. X with pmf $f(x)$

for a fixed $r \in \{1, 2, \cdots\}$

$$E(X^r) = \sum_x x^r f(x)$$

These are called as raw moments of X.

$r=1$; expectation / mean of X.

$$E(X-a)^r = \sum_x (x-a)^r f(x)$$ for some fixed $a \in \mathbb{R}$.

Central moments of X.

Take $a=0$, central moments $\Rightarrow$ raw moments.

Take $a = E(x) = \mu$

$$E(X-\mu)^r = \sum_{x \in R_x} (x-\mu)^r f(x)$$

In particular, take $r = 2$.

$E(X-\mu)^2 = E(X - E(X))^2$ = variance of X.

The positive square root of variance is called as standard deviation of X.

$$\sigma^2 = E(X-\mu)^2$$

Interpretation of $\sigma^2 = \text{var}(X)$.

Measuring squared variability of X around a.

$$= E(X-a)^2$$

We are interested in minimizing $E(X-a)^2$ wrt. a

(Approximating the r.v. X by a constant a and the error in this approximation is $E(X-a)^2$)

Observe: $E(X-a)^2 = E(X^2 - 2aX + a^2)$

$$= E(X^2) - 2a\,E(X) + a^2$$

Diff. w.r.t. a $\quad -2E(X) + 2a = 0 \Rightarrow \boxed{a = E(X) = \mu}$

Another interpretation of variance.

Given a random variable X and a number $a \in \mathbb{R}$.

$$(X-a)^2 = [(X-\mu) + (\mu-a)]^2 \quad \text{where } \mu = E(X)$$

$$= (X-\mu)^2 + (\mu-a)^2 + 2(X-\mu)(\mu-a)$$

$$E(X-a)^2 = E(X-\mu)^2 + E(\mu-a)^2 + E[2(X-\mu)(\mu-a)]$$

$$= \underbrace{E(X-\mu)^2}_{var(X)} + (\mu-a)^2 + 2(\mu-a)\underbrace{\cancel{E(X-\mu)}}_{0}$$

$$E(X-\mu) = E(X)-\mu = \mu-\mu = 0$$

$$\boxed{E(X-a)^2 = var(X) + (\mu-a)^2}$$

$$\Rightarrow \boxed{E(X-a)^2 \geq E(X-\mu)^2}$$

<u>MGF</u>.

$$M_X(t) = E(e^{tX}) = \sum_{x \in R_X} e^{tx} f(x)$$

$$E(e^{tX}) = E\left(1 + tX + \frac{t^2X^2}{2!} + \frac{t^3X^3}{3!} + \cdots\right)$$

??

$$= E(1) + E(tX) + E\left(\frac{t^2X^2}{2!}\right) + \cdots$$

$$E(e^{tX}) = 1 + t\,E(X) + \frac{t^2}{2!}E(X^2) + \frac{t^3}{3!}E(X^3) + \cdots$$

$$\frac{d}{dt}E(e^{tX}) = \underbrace{E(X) + \frac{2t}{2!}E(X^2) + \frac{3t^2}{3!}E(X^3) + \cdots}$$

$$\frac{d}{dt} E(e^{tx}) \Big|_{t=0} = E(x)$$

$$\boxed{\frac{d}{dt} M_x(t) \Big|_{t=0} = E(x)}$$

$$\frac{d^2}{dt^2} E(e^{tx}) = E(x^2) + \frac{3 \cdot 2 \cdot t}{3!} E(x^3) + \cdots$$

$$\frac{d^2}{dt^2} E(e^{tx}) \Big|_{t=0} = E(x^2)$$

$$\boxed{\frac{d^2}{dt^2} M_x(t) \Big|_{t=0} = E(x^2)}$$

$$\boxed{\frac{d^r}{dt^r} M_x(t) \Big|_{t=0} = E(x^r)}$$

Properties:

i) $M_x(t=0) = 1$

2) $M_x(t)$ as a f^n of t should be defined in a small neighbourhood around $t=0$.

$$\frac{d}{dt} M_x(t) \Big|_{t=0} = E(x)$$

Let X be a r.v. with MGF $M_X(t)$.

Find the MGF of r.v. $aX + b$ for some $a, b \in \mathbb{R}$ and $a \neq 0$.

$$M_{aX+b}(t) = E\left(e^{t(aX+b)}\right)$$

$$= E\left(e^{atX} \cdot e^{bt}\right)$$

$$= e^{bt} E\left(e^{(at)X}\right)$$

$$\boxed{M_{aX+b}(t) = e^{bt} M_X(at)}$$

Q. $X \sim \text{Binomial}(n, p)$

$$M_X(t) = ((1-p) + pe^t)^n$$

Compute mean & variance of X??

Notice: $E(X^2) = var(X) + \mu^2$

$\Rightarrow var(X) = E(X^2) - \mu^2 = E(X^2) - (E(X))^2$

$$\frac{d}{dt} M_X(t) \Big|_{t=0} = np$$

$$var(X) = np(1-p)$$

Chebyshev's inequality:

Let X be a non-negative random variable with finite expectation.

Let $t > 0$ be any positive real number.

We define a new random variable Y from X as follows.

$$Y = 0 \qquad \text{if } X < t$$

$$Y = t \qquad \text{if } X \geq t$$

$P(Y = 0) = \underbrace{P(X < t)}_{1-p}$ and $P(Y = t) = \underbrace{P(X \geq t)}_{p}$

$$E(Y) = 0 \cdot P(Y=0) + t\,P(Y=t)$$

$$= t\,P(X \geq t)$$

Note: $X \geq Y$

$$\Rightarrow E(X) \geq E(Y) = t\,P(X \geq t)$$

$$\Rightarrow \boxed{P(\underbrace{X \geq t}) \leq \boxed{\frac{E(X)}{t}}} \quad \text{--- } (*)$$

<u>Chebyshev's inequality:</u>

Let X be a r.v. with mean μ and variance σ^2. For any real no. $t > 0$

$$P\left((X-\mu)^2 \geq t^2\right) \leq \frac{E(X-\mu)^2}{t^2} \quad \text{... from } (*)$$

Notice: i) $E(X-\mu)^2 = \sigma^2 = var(X)$

ii) the events $\boxed{|X-\mu| \geq t}$ and $\boxed{(X-\mu)^2 \geq t^2}$

are same !!

$$\boxed{P(|X-\mu| \geq t) \leq \frac{\sigma^2}{t^2}} \quad (*)$$

Expectation of a continuous r.v.

Let X be a continuous r.v. with pdf $f_X(x)$. The expectation of X is defined as

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\,dx \quad \text{if it exists.}$$

$$\left[E(X) \text{ exists if } \int_{-\infty}^{\infty} |x| f_X(x)\,dx < \infty \right]$$

Ex: $X \sim U(a, b)$

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\, dx = \int_a^b x \frac{1}{b-a} dx$$

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & \text{o.w.} \end{cases}$$

$$E(X) = \frac{x^2}{2} \cdot \frac{1}{b-a} \Big]_a^b = \frac{a+b}{2}$$

Ex: $X \sim \text{gamma}(\alpha, \lambda)$

$$f_X(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \quad \text{for } x > 0$$

$$= 0 \quad \text{otherwise}$$

$$E(X) = \int_0^\infty x \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \, dx$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^\alpha e^{-\lambda x} \, dx = \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+1)}{\lambda^{\alpha+1}}$$

$$E(X) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\alpha \Gamma(\alpha)}{\lambda^{\alpha+1}} = \frac{\alpha}{\lambda}$$

$$\boxed{E(X) = \frac{\alpha}{\lambda}}.$$

Ex: $X \sim \text{gamma}(1, \lambda)$ $\quad \lambda > 0$

$$f_X(x) = \lambda e^{-\lambda x} \qquad x > 0$$

$$= 0 \qquad \text{o.w.}$$

$X \sim \exp(\lambda)$

$$\boxed{E(X) = \frac{1}{\lambda}}$$

Ex: Cauchy density:

$$f_X(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2} \qquad -\infty < x < \infty$$

$$E(|X|) = \int_{-\infty}^{\infty} \frac{|x|}{\pi(1+x^2)}\, dx = \int_0^{\infty} \frac{2x}{\pi(x^2+1)}\, dx$$ does NOT converge.

$\Rightarrow$ Cauchy density does not have mean.

Moments of a cts. r.v.

i) $E(X^m) = \int_{-\infty}^{\infty} x^m f_X(x)\,dx$: m^{th} raw moment.

ii) $E(X-a)^m = \int_{-\infty}^{\infty} (x-a)^m f_X(x)\,dx$: m^{th} central moment.

variance of $X = \sigma^2 = E(X-\mu)^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)\,dx$

where $\mu = E(X)$

$\underline{\text{Ex:}}$ $X \sim \text{gamma}(\alpha, \lambda)$

$$E(X^m) = \int_0^\infty x^m \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \, dx$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{m+\alpha-1} e^{-\lambda x} \, dx$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+m)}{\lambda^{\alpha+m}}$$

$$\boxed{E(X^m) = \frac{\alpha(\alpha+1)\cdots(\alpha+m-1)}{\lambda^m}}$$

variance of X

$= E(X^2) - (E(X))^2$

$= \frac{\alpha(\alpha+1)}{\lambda^2} - \frac{\alpha^2}{\lambda^2}$

$= \boxed{\frac{\alpha}{\lambda^2}}$

In particular,

$$X \sim \exp(\lambda) = \text{gamma}(1,\lambda)$$

$$E(X^m) = \frac{m!}{\lambda^m}$$

$$\text{var}(X) = \frac{1}{\lambda^2}$$

Recall symmetric densities and symmetric r.v.s

Let X be a symmetric r.v.

$\Rightarrow$ X and $-X$ have same density.

for any $m \in \mathbb{N}$

X^m and $(-X)^m$ have same density.

Let $m \in \mathbb{N}$ be an odd number.

$$(-x)^m = -x^m$$

Then x^m and $-x^m$ have same density.

$$\Rightarrow E(x^m) = E(-x^m) = -E(x^m)$$ (s.t. condition that expectations exist.)

$$\Rightarrow E(x^m) = 0$$

$\Rightarrow$ Any odd ordered raw moment of symmetric density is zero. Ex: std. normal.

MGF:

$$M_X(t) = E(e^{tX}) \quad \text{if it exists.}$$

Ex: $X \sim N(\mu, \sigma^2)$

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \quad -\infty < x < \infty$$

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \, dx$$

substitute $y = x - \mu$

$$M_X(t) = \int_{-\infty}^{\infty} e^{t(y+\mu)} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{y^2}{2\sigma^2}} \, dy$$

$$M_X(t) = e^{\mu t} \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{\left(ty - \frac{y^2}{2\sigma^2}\right)} dy$$

<u>Notice:</u> $$ty - \frac{y^2}{2\sigma^2} = -\frac{(y-\sigma^2 t)^2}{2\sigma^2} + \frac{\sigma^2 t^2}{2}$$

$$M_X(t) = e^{\mu t} \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\sigma^2 t)^2/2\sigma^2} e^{\frac{\sigma^2 t^2}{2}} dy$$

$$= e^{\mu t + \frac{1}{2}\sigma^2 t^2} \int_{-\infty}^{\infty} \underbrace{\boxed{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\sigma^2 t)^2}{2\sigma^2}}}}_{\text{pdf of } N(\sigma^2 t,\, \sigma^2)} dy$$

$$\boxed{M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}}$$

$$E(X) = \frac{d}{dt} M_X(t) \Big|_{t=0} = \frac{d}{dt} e^{\mu t + \sigma^2 t^2/2} \Big|_{t=0}$$

$$\boxed{E(X) = \mu}$$

$$E(X^2) = \sigma^2 + \mu^2 \quad \Rightarrow \quad \boxed{\mathrm{var}(X) = \sigma^2}$$

In general, if $X \sim \text{Normal}(\text{mean}, \text{variance})$

$$M_X(t) = e^{(\text{mean})t + \frac{1}{2}(\text{variance})t^2}$$

$$X \sim N(0,1) \qquad \boxed{M_X(t) = e^{t^2/2}}$$

Ex: $X \sim \text{gamma}(\alpha, \lambda)$

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

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

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha)}{(\lambda-t)^\alpha}$$

$$\boxed{M_X(t) = \left(\frac{\lambda}{\lambda-t}\right)^\alpha}$$

$$-\infty < t < \lambda$$

Quartiles:

A point $q_1 \in \mathbb{R}$ is called as 1^{st} quartile if

$$P(X \leq q_1) = \frac{1}{4}$$

$$F_X(q_1) = \frac{1}{4}$$

A point $m \in \mathbb{R}$ is called 2^{nd} quartile or median if

$$P(X \leq m) = \frac{1}{2}$$

$$F_X(m) = \frac{1}{2}$$

A point is called as third quartile if

$$P(X \leq q_3) = \frac{3}{4}$$

$$F_X(q_3) = \frac{3}{4}$$

<u>Mode</u>:

Let X be a r.v. with density $f_X(x)$.

$$M = \arg\max_{x} f_X(x)$$

M is called as <u>mode</u> of X.

Ex: $X \sim \text{Binomial}(3, \frac{1}{2})$

compute mode, median

x	0	1	2	3
$p(x)$	$1/8$	$3/8$	$3/8$	$1/8$

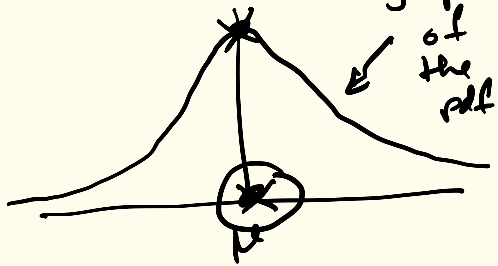
mean $= np = \frac{3}{2}$

mode $= 1, 2$

median $= 1$

Ex: $X \sim N(\mu, \sigma^2)$

$\text{mean}(X) = \mu = E(X)$



Let m be the median.

$$F_X(m) = 0.5$$

$$\Rightarrow P(X \le m) = 0.5$$

$$\Rightarrow P\left(\frac{X-\mu}{\sigma} \le \frac{m-\mu}{\sigma}\right) = \frac{1}{2}$$

Φ: CDF of std. normal.

$$\Rightarrow \Phi\left(\frac{m-\mu}{\sigma}\right) = \frac{1}{2}$$

$$\Rightarrow \frac{m-\mu}{\sigma} = 0$$

$$\Rightarrow \boxed{m = \mu}$$

In case $N(\mu, \sigma^2)$
mean = median = mode

Ex: $X \sim \text{uniform}(a,b)$

$$E(X) = \mu = \frac{a+b}{2}$$

Let m be the median

$$F(m) = \frac{1}{2} \Rightarrow \frac{m-a}{b-a} = \frac{1}{2} \Rightarrow \boxed{m = \frac{a+b}{2}}$$

Ex: $X \sim \exp(\lambda)$

$F(m) = \frac{1}{2}$ where m: median

$$1 - e^{-\lambda m} = \frac{1}{2} \Rightarrow e^{-\lambda m} = \frac{1}{2} \Rightarrow -\lambda m = \log \frac{1}{2}$$

$$\Rightarrow \boxed{m = -\frac{1}{\lambda} \log \frac{1}{2}}$$