Moment Generating Functions

Moments

Single function that determines all

Moment generating function (MGF)

Definition

Examples

Properties

EX_MACHINA

Moments

Expectations of powers of X are called Moments

$E(X)$ $E(X^2)$ $E(X^3)$

Sometimes called raw moments to distinguish from central moments

$E(X-\mu)^n$

Determine mean, variance, many other distribution properties

General method to determine all moments

Moment Generating Function (MGF)

Maps a random variable X to a real function M

M: $\mathbb{R} \to \mathbb{R}$

$$M(t) \stackrel{\text{def}}{=} M_X(t) \stackrel{\text{def}}{=} E[e^{tX}] = \begin{cases} \sum p(x)\, e^{tx} & \text{discrete} \\ \int f(x)\, e^{tx}\, dx & \text{continuous} \end{cases}$$

Determined by distribution p or f

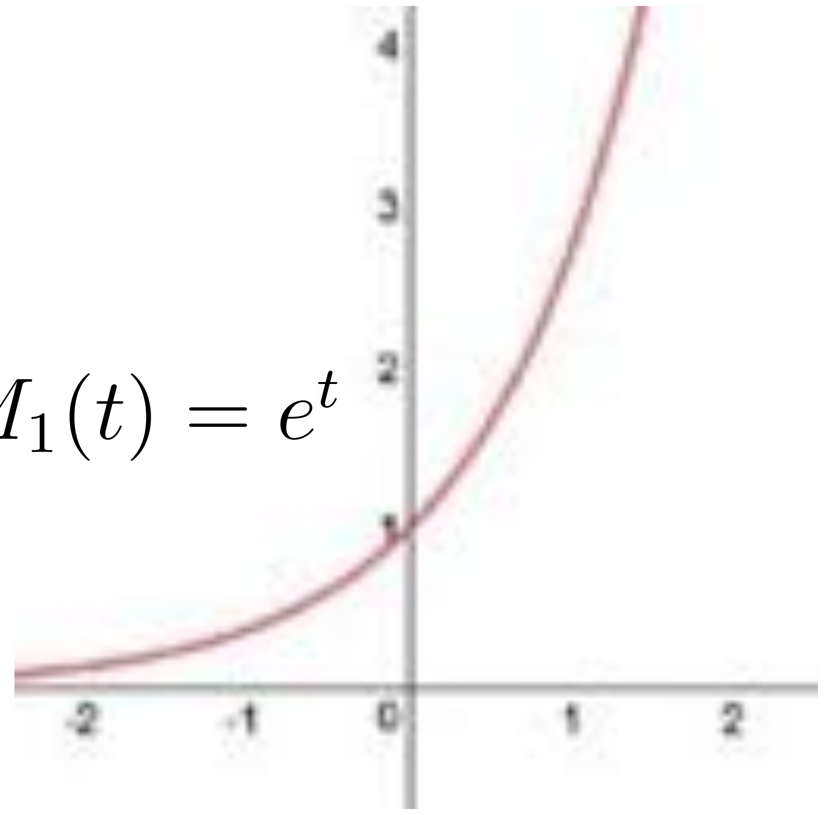
X more convenient

Discrete

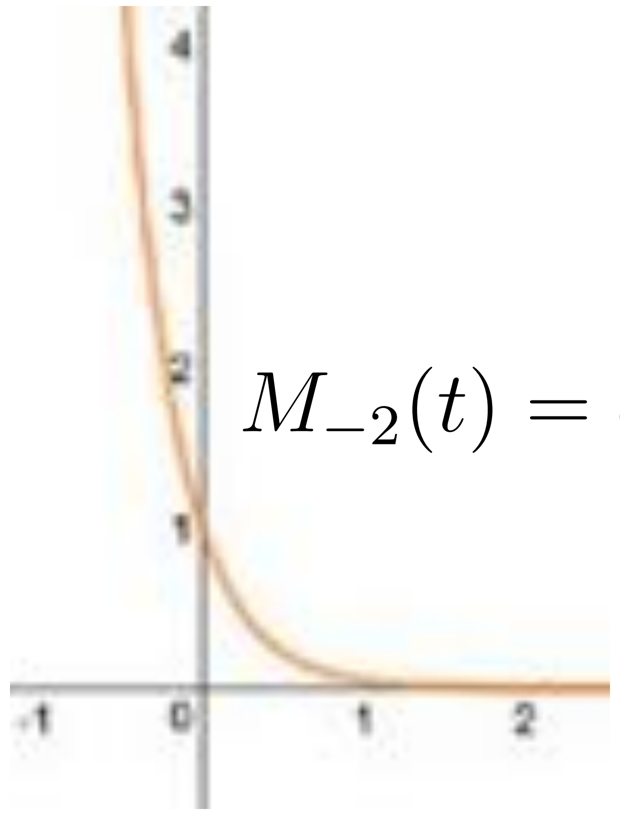
Continuous

Few Values

$M_1(t) = e^t$



$M_{-2}(t) = e^{-2t}$

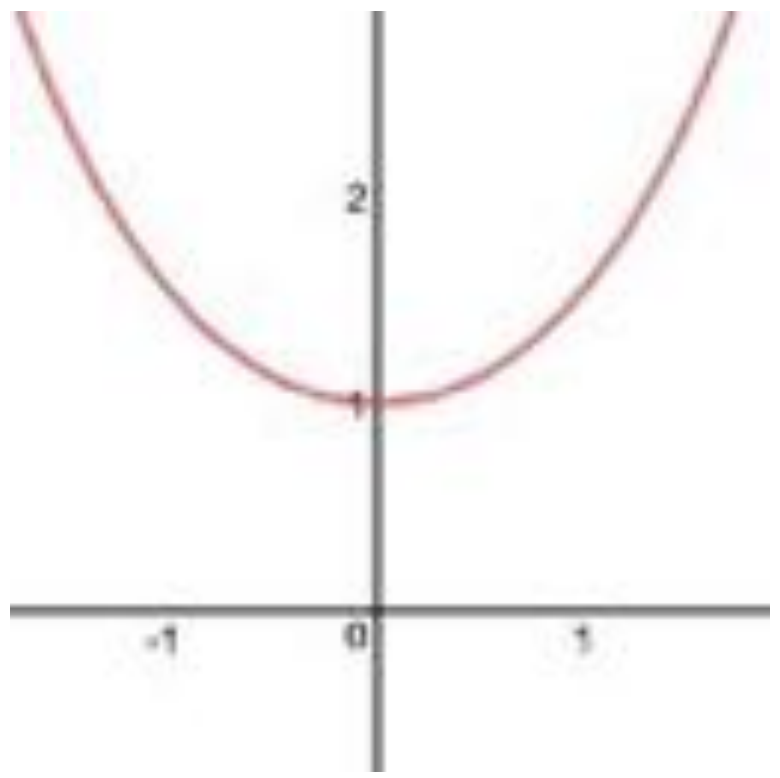


One value (constant)

$$X = c \qquad M_c(t) = E(e^{tX}) = \sum p(x)\, e^{tx} = e^{ct}$$

$$p_{-1} = p_1 = \frac{1}{2}$$

$M(t) = \frac{1}{2}e^{-t} + \frac{1}{2}e^t$



Two arbitrary values

$$p(c_1) = p_1 \quad p(c_2) = p_2 \qquad M(t) = E(e^{tX}) = p_1\, e^{c_1 t} + p_2\, e^{c_2 t}$$

0, 1 values (Bernoulli)

$X \sim B_p$ $\qquad$ $p_0 = 1 - p$ $\qquad$ $p_1 = p$

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

Properties

Positive

$$M(t) = E(e^{tX}) > 0$$

$0 \rightarrow 1$

$$M(0) = E(e^{0X}) = E(e^0) = 1$$

Translation and Scaling

Translation X → X+b

$$M_{X+b}(t) = E(e^{t(X+b)}) = E(e^{tX} \cdot e^{tb}) = e^{tb} \cdot E(e^{tX}) = e^{tb} M_X(t)$$

Scaling X → aX

$$M_{aX}(t) = E(e^{t(aX)}) = E(e^{atX}) = M_X(at)$$

Translation and Scaling X → aX+b

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

$$M_c(t) = e^{ct} \longrightarrow M_{a \cdot c+b} = e^{bt} \cdot M_c(at) = e^{bt} \cdot e^{cat} = e^{(ac+b) \cdot t}$$

Addition

Independent variables

The MGF of the sum is the product of the MGF's

Two variables

X $\perp\!\!\!\perp$ Y

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$$

$$M_{X+Y}(t) = E[e^{t(X+Y)}] = E[e^{tX} \cdot e^{tY}] = E[e^{tX}] \cdot E[e^{tY}] = M_X(t) \cdot M_Y(t)$$

n variables

$X_1, X_2, \ldots, X_n \quad \perp\!\!\!\perp$

$$X \stackrel{\text{def}}{=} X_1 + X_2 + \ldots + X_n$$

$$M_X(t) = \prod_{i=1}^n M_{X_i}(t)$$

Average

$X_1, X_2, \ldots, X_n \quad \perp\!\!\!\perp$

$$\overline{X} \stackrel{\text{def}}{=} \frac{X_1+X_2+\ldots+X_n}{n}$$

$$M_{\overline{X}}(t) = \prod_{i=1}^n M_{X_i}\left(\frac{t}{n}\right)$$

Moment Generation

$E(X^n)$ — n^{th} "raw" moment of X

$M_X(t)$ determines ("generates") all moments of X

$$E(X) = M'_X(0) \qquad E(X^2) = M''_X(0) \qquad E(X^n) = M_X^{(n)}(0)$$

$$e^y = 1 + \frac{y}{1!} + \frac{y^2}{2!} + \frac{y^3}{3!} + \dots$$

$$\begin{aligned} M_X(t) = E(e^{tX}) &= E\left(1 + \frac{t}{1!}X + \frac{t^2}{2!}X^2 + \frac{t^3}{3!}X^3 + \dots\right) \\ &= 1 + \frac{t}{1!}E(X) + \frac{t^2}{2!}E(X^2) + \frac{t^3}{3!}E(X^3) + \dots \end{aligned}$$

Moment Generating Functions

Moments

Single function that determines all

Moment generating function (MGF)

Definition

Examples

Properties

EX_MACHINA


iStock
iStock
iStock
iStock
iStock
iStock
iStock
Conveyor belt

Binomial

$B_{p,n} \quad 0 \le p \le 1, \; n \ge 0$

$p(k) = \binom{n}{k} p^k (1-p)^{n-k}$

$M(t) = \; [1 + p(e^t - 1)]^n$

$$M(t) = E(e^{tX}) = \sum_{k=0}^{n} p(k) \cdot e^{tk}$$

$$= \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \cdot e^{tk}$$

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

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

$$= [1 + p(e^t - 1)]^n$$

Use….. Binomial Thm

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k \, b^{n-k}$$

Moments

$E(X)$

$$M'(t) = n(pe^t + 1 - p)^{n-1} \cdot pe^t$$

$$M'(0) = n(pe^0 + 1 - p)^{n-1} \cdot pe^0 = np$$

Poisson

$P_\lambda, \ \lambda > 0$

$p(k) = e^{-\lambda} \frac{\lambda^k}{k!}$

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

$$M(t) \quad = \ E(e^{tX}) \ = \ \sum_{k=0}^{\infty} p(k) \cdot e^{tk}$$

$$= \ \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \cdot e^{tk}$$

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

$$= \ e^{\lambda e^t - \lambda} \cdot \sum_{k=0}^{\infty} e^{-\lambda e^t} \frac{(\lambda e^t)^k}{k!} \quad = \ e^{\lambda(e^t - 1)}$$

Standard Normal

$\mathcal{N}(0,1)$ $\quad f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ $\quad M(t) = e^{\frac{t^2}{2}}$

$$M(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx$$

$$\int_{-\infty}^{\infty} \to \int$$

$$= \int e^{tx} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int \exp[-\frac{x^2-2tx}{2}]\, dx$$

$$= \frac{1}{\sqrt{2\pi}} \int \exp[-\frac{(x-t)^2}{2} + \frac{t^2}{2}]\, dx$$

$$= e^{\frac{t^2}{2}} \frac{1}{\sqrt{2\pi}} \int \exp[-\frac{(x-t)^2}{2}]\, dx$$

$$= e^{\frac{t^2}{2}}$$

“Abnormal” Normal

Standard normal $Z \sim \mathcal{N}(0,1)$ $M_Z(t) = e^{\frac{t^2}{2}}$

General normal $X \sim \mathcal{N}(\mu, \sigma^2)$ $M_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$

$X = \sigma Z + \mu$

$$M_X(t) = M_{aZ+b}(t) = e^{bt} M_Z(at)$$

$$= e^{\mu t} \cdot e^{\frac{\sigma^2 t^2}{2}}$$

$$= e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

Properties

Informally

If $M_X(t) = M_Y(t)$ then X and Y have the same distribution

Can invert $M_X(t)$ to obtain X

If $M_{X_n}(t) \rightarrow M_X(t)$ then the $f_{X_n} \rightarrow f_X$

Poisson and Binomial

Binomial	Poisson	Convergence
$B_{p,n}(k) = \binom{n}{k} p^k (1-p)^{n-k}$	$P_\lambda(k) = e^{-\lambda} \frac{\lambda^k}{k!}$	$B_{\frac{\lambda}{n},n} \to P_\lambda$
$M_{p,n}(t) = [1 + p(e^t - 1)]^n$	$M_\lambda(t) = e^{\lambda(e^t-1)}$	$M_{\frac{\lambda}{n},n}(t) \to M_\lambda(t)$?

$(1 + \frac{1}{n})^n \to e$

$(1 + \frac{a}{n})^n = [(1 + \frac{a}{n})^{\frac{n}{a}}]^a \to e^a$

$$M_{\frac{\lambda}{n},n}(t) = [1 + \frac{\lambda}{n}(e^t - 1)]^n \to e^{\lambda(e^t-1)} = M_\lambda(t)$$

Moment Generating Functions

Moments

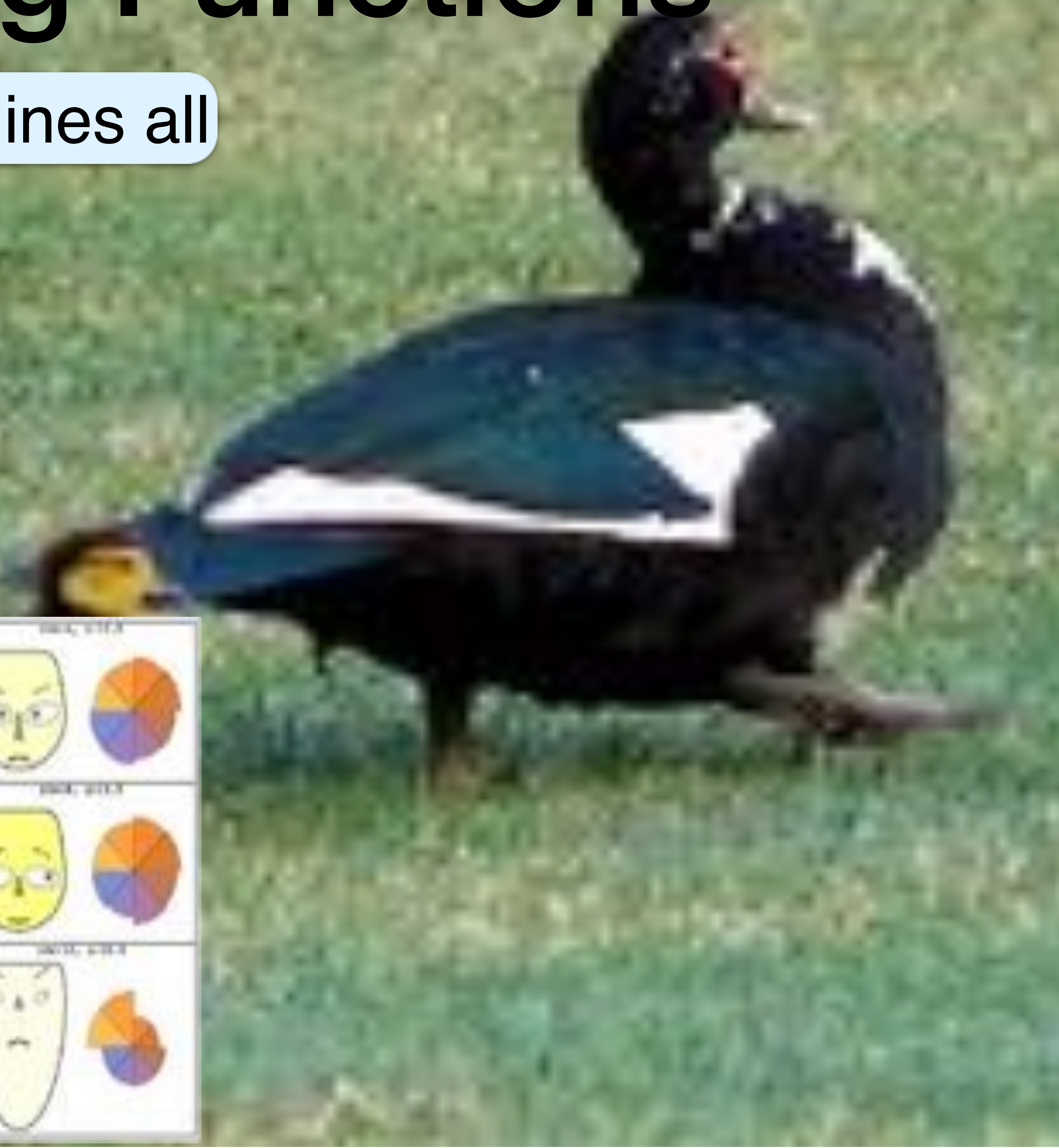
Single function that determines all

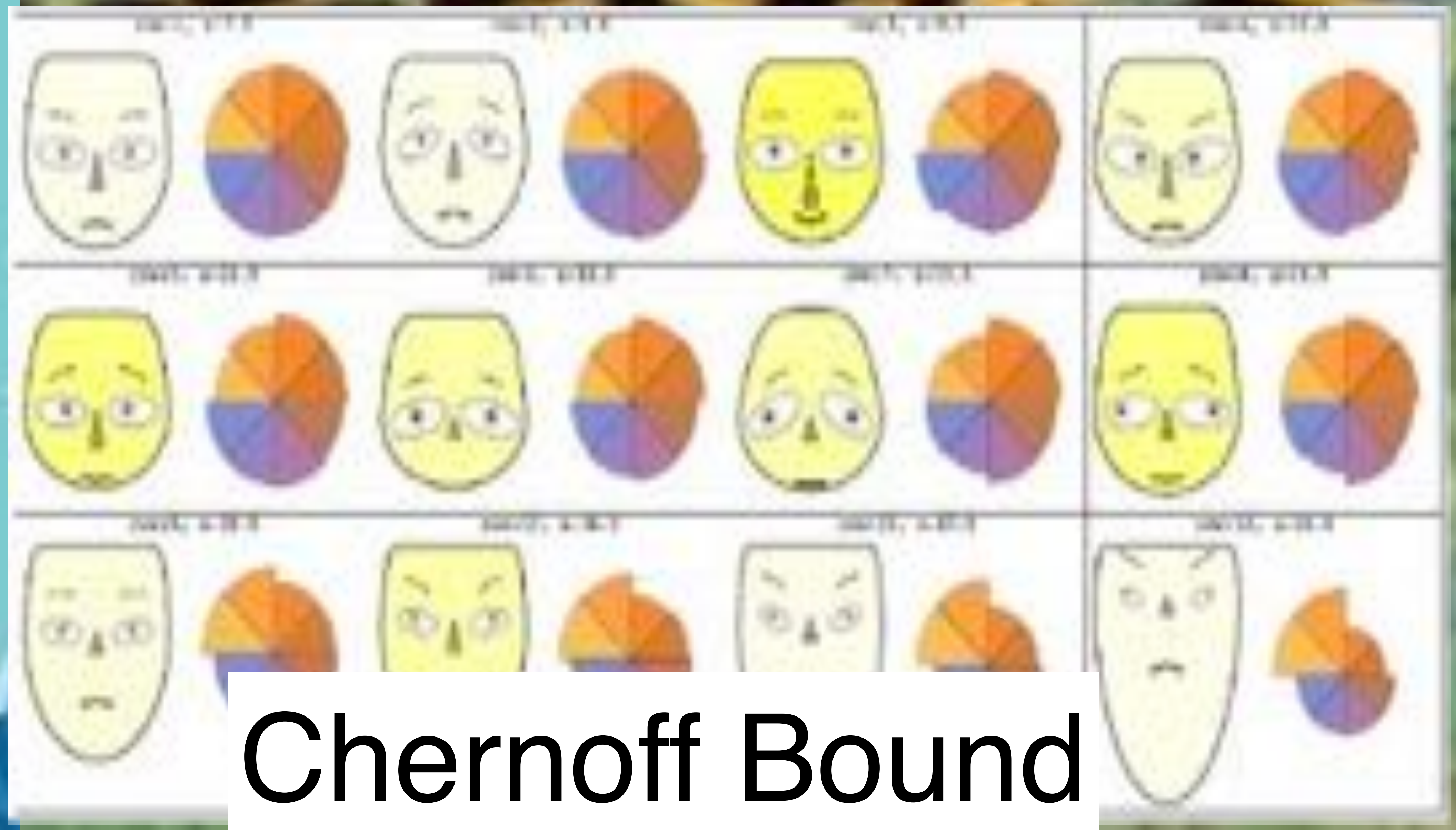
Moment generating function (MGF)

Definition

Examples

Properties





Chernoff Bound

Chernoff Bound

Herman Chernoff

1923 -

Statistician

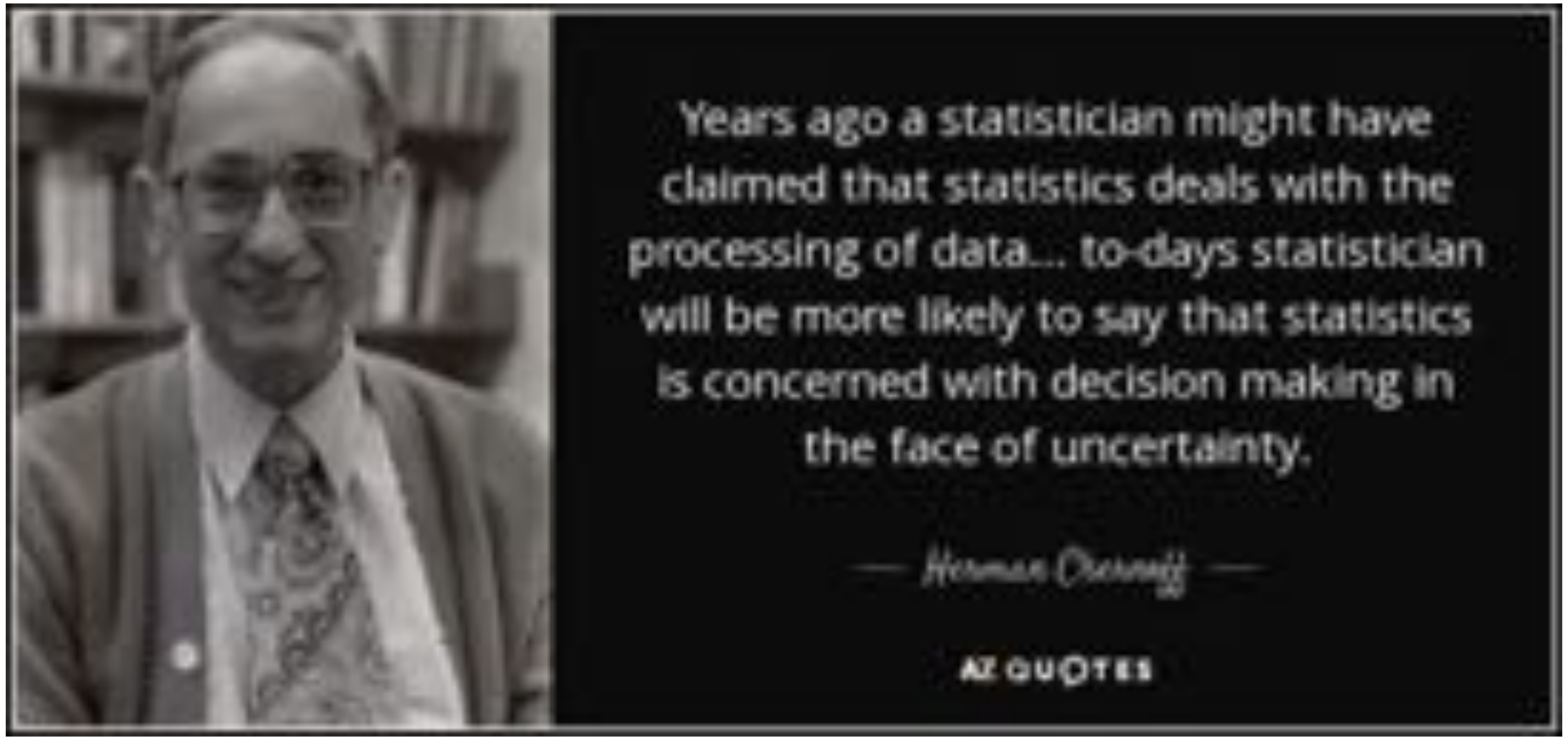
UIUC, Stanford, MIT, Harvard

Broad view of statistics

Chernoff bound

Markov → Chebyshev → Chernov

Chernoff faces



The New York Times
April 1, 2008

Chernoff Bound

$X \sim B_{p,n}$

$\mu = pn$

$P\left(X \geq (1+\delta)\mu\right)$

Decrease

Markov - constant

Chebyshev - linear

Chernoff - exponential

$\forall a \quad \forall t \geq 0$

$X \geq a \quad \leftrightarrow \quad tX \geq ta \quad \leftrightarrow \quad e^{tX} \geq e^{ta}$

$$P(X \geq a) = P(e^{tX} \geq e^{ta}) \overset{\text{Markov}}{\leq} \frac{E(e^{tX})}{e^{ta}}$$

$E(e^{tX})$

Evaluate

Bound

Incorporate

Simplify

Evaluate $E(e^{tX})$

Same as moment generating functions

Self-contained

$X \sim B_{p,n}$

$X = \sum_{i=1}^{n} X_i$

$X_i \sim B_p \quad \perp\!\!\!\perp$

$$E\left(e^{tX}\right) = E(e^{t\Sigma X_i}) = E(e^{\Sigma t X_i}) = E\left(\prod_{i=1}^{n} e^{tX_i}\right) \overset{\perp\!\!\!\perp}{=} \prod_{i=1}^{n} E\left(e^{tX_i}\right) = [(1-p) + pe^t]^n$$

$$E(e^{tX_i}) = P(X_i = 0) \cdot e^{t \cdot 0} + P(X_i = 1) \cdot e^{t \cdot 1} = (1-p) + pe^t$$

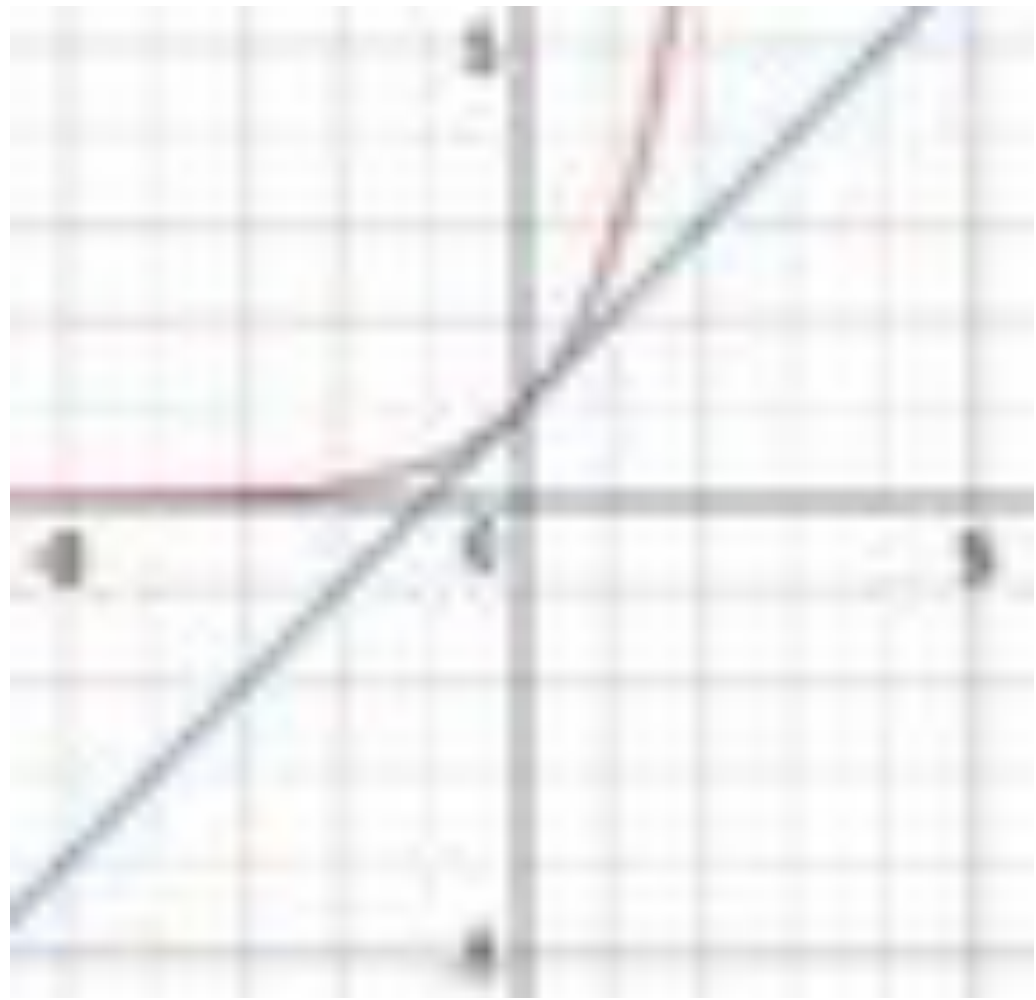
$$E\left(e^{tX}\right) = [(1-p) + pe^t]^n$$

Bound $E(e^{tX})$

$$E\left(e^{tX}\right) = [(1-p)+pe^t]^n \leq \left(e^{p(e^t-1)}\right)^n = e^{np(e^t-1)} = e^{\mu(e^t-1)}$$

$$(1-p)+pe^t = 1+p(e^t-1) \leq e^{p(e^t-1)}$$

$$1+x \leq e^x$$



$$e^x = 1+x+\frac{x^2}{2}+\ldots \geq 1+x$$

$$E\left(e^{tX}\right) \leq e^{\mu(e^t-1)}$$

Incorporate in Markov ≤

$X \sim B_{p,n}$

$\forall a \quad \forall t \geq 0$

$P(X \geq a) = P(e^{tX} \geq e^{ta}) \leq \frac{E(e^{tX})}{e^{ta}}$

$E\left(e^{tX}\right) \leq e^{\mu(e^t-1)}$

$P(X \geq a) \leq \frac{e^{\mu(e^t-1)}}{e^{ta}}$

$a = (1+\delta)\mu$

$\delta \geq 0$

$\forall t \geq 0$

$P(X \geq (1+\delta)\mu) \leq \frac{e^{\mu(e^t-1)}}{e^{\mu(1+\delta)t}} = e^{\mu((e^t-1)-t(1+\delta))}$

Optimize t

Optimize t

$\forall t \geq 0$

$$P\left(X \geq (1+\delta)\mu\right) \leq e^{\mu((e^t-1)-t(1+\delta))}$$

Find t minimizing

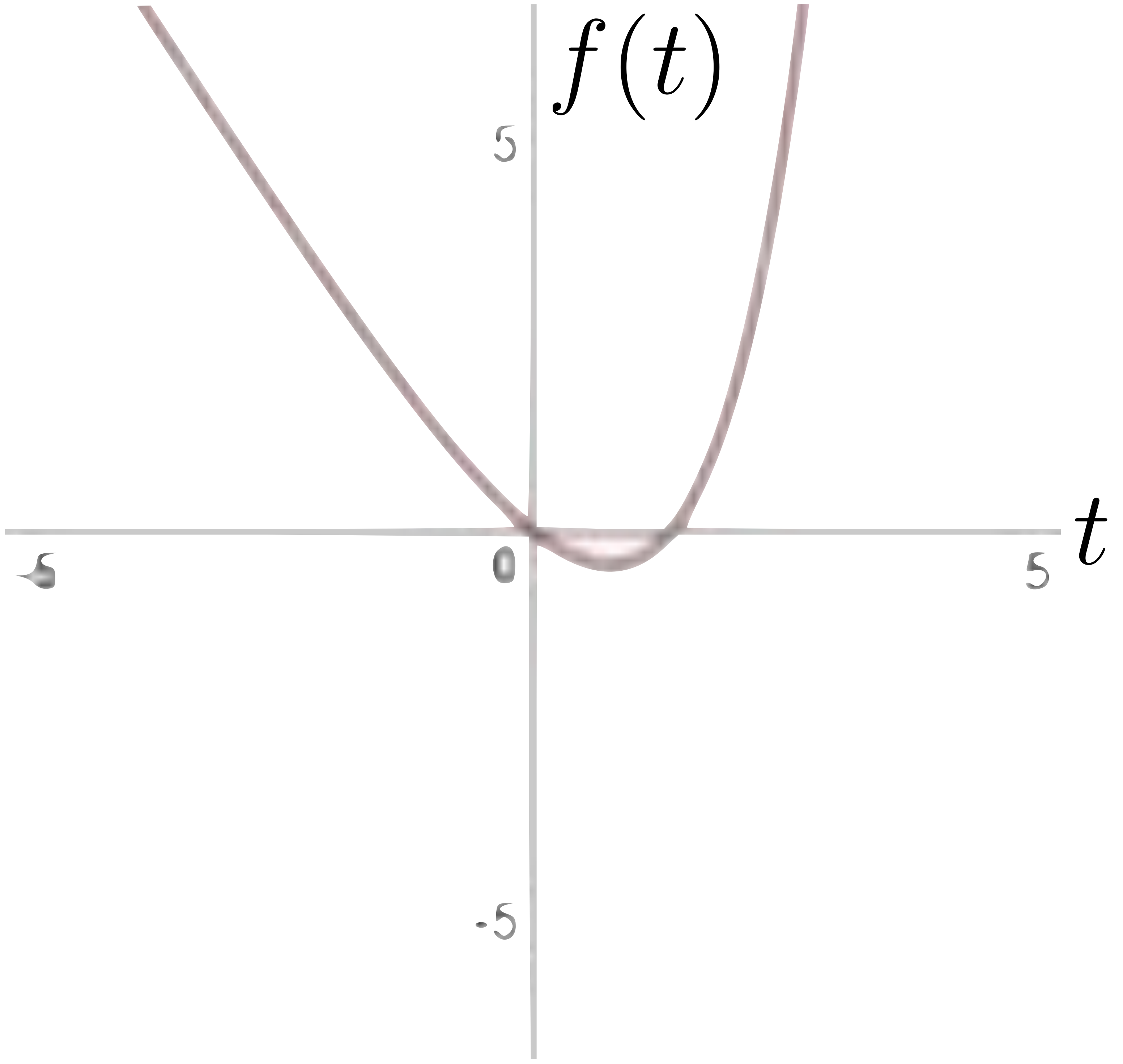
$$f(t) \stackrel{\text{def}}{=} (e^t - 1) - t(1+\delta)$$

$$f'(t) = e^t - (1+\delta) = 0$$

$$e^t = 1+\delta \qquad t = \ln(1+\delta)$$

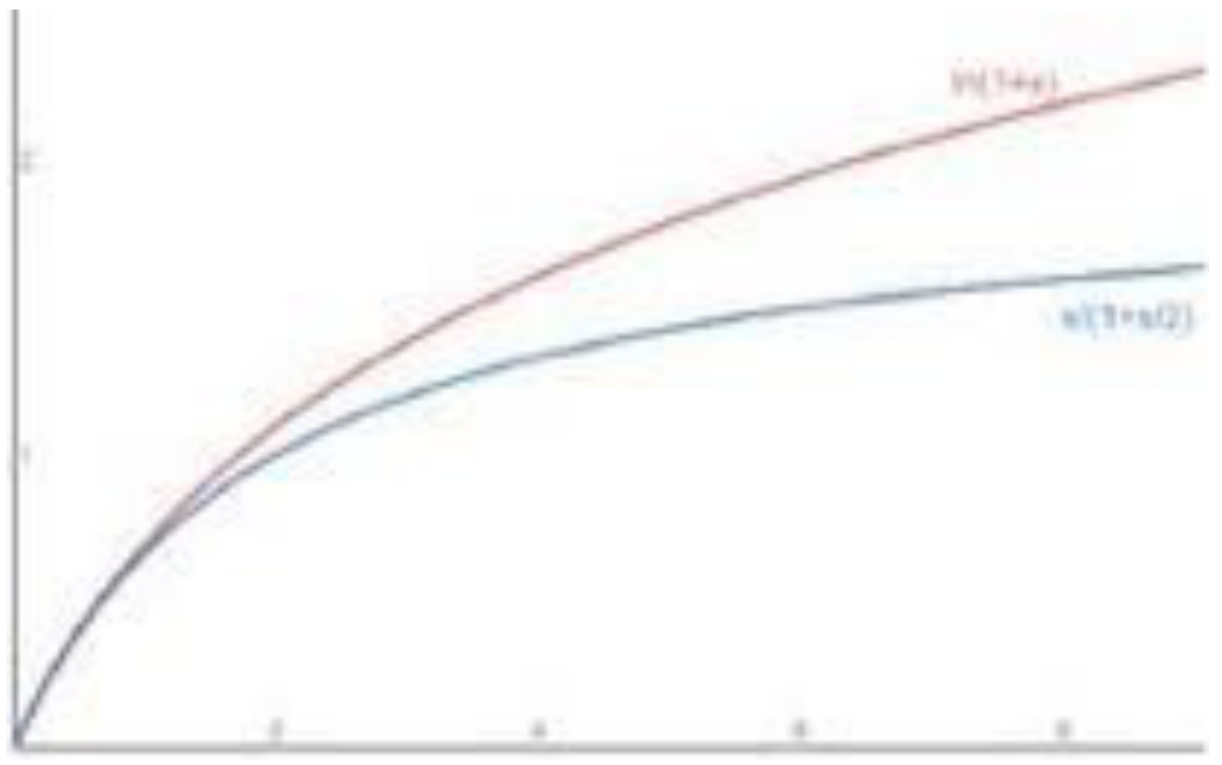
$$f''(t) = e^t \geq 0$$

$$P\left(X \geq (1+\delta)\mu\right) \leq e^{\mu((e^t-1)-t(1+\delta))} = e^{\mu[\delta-(1+\delta)\ln(1+\delta)]}$$



Final Simplification

$P\left(X \geq (1+\delta)\mu\right) \leq e^{\mu[\delta-(1+\delta)\ln(1+\delta)]}$ $\leq \quad e^{-\frac{\delta^2}{2+\delta}\mu}$



Inequality $\ln\left(1+x\right) \geq \frac{x}{1+\frac{x}{2}} \qquad \forall x \geq 0$

Define $f(x) = \ln(1+x) - \frac{x}{1+\frac{x}{2}}$ Show $f(x) \geq 0 \qquad \forall x \geq 0$

$f(0) \;=\; \ln 1 - \frac{0}{1} \;=\; 0$

$f'(x) = \frac{1}{1+x} - \frac{(1+\frac{x}{2})-\frac{x}{2}}{(1+\frac{x}{2})^2} = \frac{1}{1+x} - \frac{1}{(1+\frac{x}{2})^2} = \frac{1}{1+x} - \frac{1}{1+x+\frac{x^2}{2}} \geq 0$

$f(x) \geq 0 \qquad \forall x \geq 0$

$\delta - (1+\delta)\ln(1+\delta) \leq \delta - (1+\delta)\frac{\delta}{1+\frac{\delta}{2}} = \delta\left[1 - \frac{1+\delta}{1+\frac{\delta}{2}}\right] = \delta[\frac{2+\delta-2-2\delta}{2+\delta}] = \frac{-\delta^2}{2+\delta}$

Chernoff Bound

$X \sim B_{p,n}$

$\delta \geq 0$

$$P\left(X \geq (1+\delta)\mu\right) \leq \; e^{-\frac{\delta^2}{2+\delta}\mu}$$

$$P\left(X \leq (1-\delta)\mu\right) \leq \; e^{-\frac{\delta^2}{2}\mu}$$

Poll

47% vote D

6000 samples

X : # vote D

$X \sim B_{0.47,6000}$

Wrong:

$X \geq 0.5 \cdot 6{,}000 = 3{,}000$

$$\mu = np = 6000 \cdot 0.47 = 2820$$

$$3000 = (1+\delta)\mu$$

$$1 + \delta = \frac{3000}{2820} = \frac{6000 \cdot 0.5}{6000 \cdot 0.47} = \frac{50}{47} \approx 1.06383$$

$$\delta \approx 0.0638$$

P(wrong)

$$P(X \geq 3000) = P(X \geq (1+\delta)\mu) \leq e^{-\frac{\delta^2}{2+\delta}\mu} \approx 0.38\%$$

Chernoff Bound

Exponential bound on exceeding mean by a constant factor

Next

Central Limit Theorem