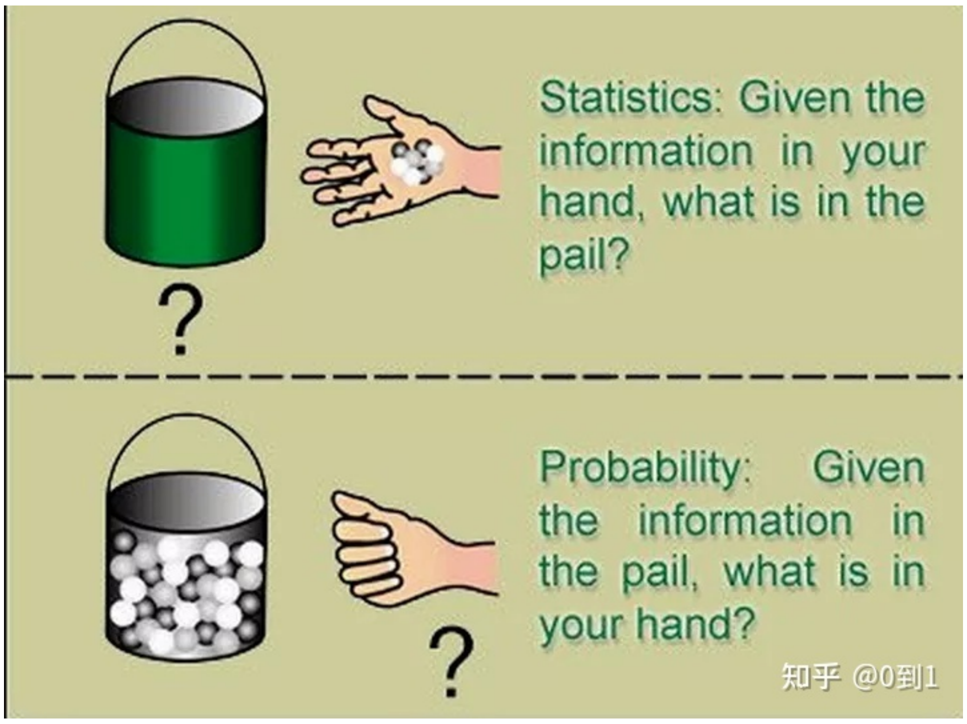
Chapter 23.

Probability and Statistics.



Probability

sample space S

$\boxed{\text{mutually exclusive events.}}$

$$P(x_i) = \lim_{n \to \infty} \frac{\text{number of } x_i \text{ occurs}}{\text{total trails.}}$$

Theoretical probability.

$$P(x_i) := \frac{\text{number of outcomes } x_i}{\text{total events.}}$$

example

1. tossing coins.
2. sand

Axioms

1. $0 \leq P \leq 1$
2. $\sum_i P_i = 1$
3. $A, B.$

$$\boxed{A \quad \cancel{B} \qquad \cancel{A} \quad B \qquad \underline{A+B}} = P(A) + P(B)$$

Conditional Probability

$$P(B|A)$$

$$P(A,B) = P(A) \cdot P(B|A)$$

$$P(B|A) = \frac{P(A,B)}{P(A)} = \frac{n(A,B)}{n(A)}.$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

If A, B are independent,

$$P(A,B) = P(A) \cdot P(B).$$

Bayes' theorem.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

example.

morbidity 0.5%. accuracy 99%

A: sick.

B: diagnosed.

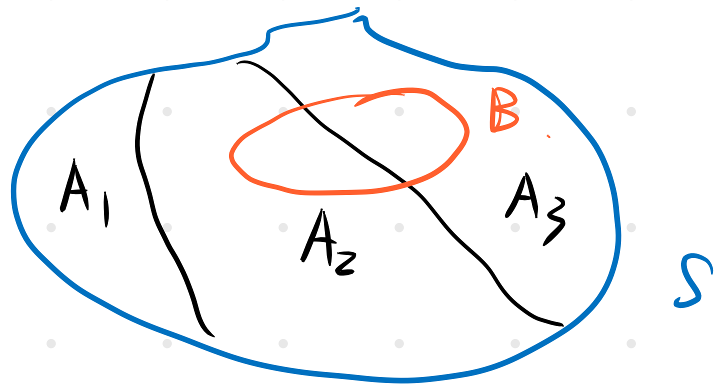
$$\frac{0.005 \times 0.99 \times 0.005}{0.995 \times 0.01 + 0.005 \times 0.99.} = 33.2\%$$

more general theorem.

$A_i \qquad \bigcup_i A_i = S$

$\forall B \subset S$.

$$P(B) = \sum_i P(A_i) P(B|A_i)$$



Permutations and Combinations.

$$P_m^n \qquad A_m^n = \frac{m!}{(m-n)!}$$

$$C_m^n \qquad \binom{m}{n} = \frac{m!}{n!\,(m-n)!} \leftarrow C_m^n = \frac{m!}{n!\,(m-n)!}$$

$$B(n_1 \cdots n_i) = \frac{n!}{n_1!\, n_2! \cdots n_i!}$$

$n \qquad i \qquad \dfrac{n!}{n_1!\, n_2! \cdots n_i!}$

multinomial coefficient.

$$(x_1 + x_2 + \cdots + x_m)^n \to B(\quad)\, \underline{x_1^{n_1}}\, x_2^{n_2} \cdots x_m^{n_m}$$

Random Variables.

Mean, deviation, correlation

probability density

$$P(x \le X \le x + dx) = f(x)\,dx$$

$$f(x) \geq 0 . \qquad \int f(x) = 1 .$$

example. 1s electron

$$|\psi| = N e^{-\frac{r}{a}}$$

$$|\psi|^2 = N^2 e^{-\frac{2r}{a}}$$

$$\int |\psi|^2 d^3 r = 4\pi N^2 \int e^{-\frac{2r}{a}} r^2 dr$$

$$= \pi a^3 N^2 = 1 \Rightarrow N = \frac{1}{\sqrt{\pi a^3}} .$$

Mean and Variance.

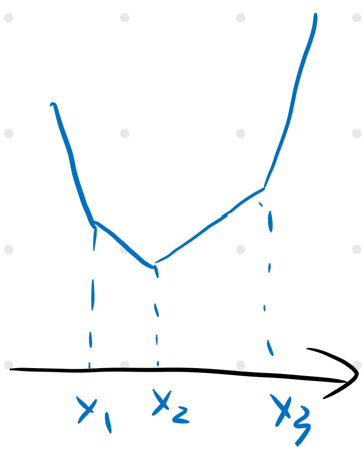
$$\bar{x} = \frac{1}{n} \sum_j x_j$$

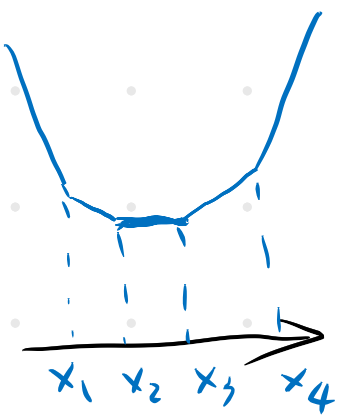
$$\langle X \rangle = \sum_i x_i \underline{p_i} = \int x \underline{f(x)\,dx}$$

$$x_g = (x_1 \cdots x_n)^{1/n}$$

$$\frac{1}{x_h} = \frac{1}{n}\left(\frac{1}{x_1} + \cdots + \frac{1}{x_n} \right).$$

$$\boxed{\sum |x_i - x|} \leftarrow \hat{x} \qquad x_2 .$$





$$\text{minimize} \sum_i (x - x_i)^2$$

$$\frac{\partial \sum_i (x - x_i)^2}{\partial x} = 2 \sum_i (x - x_i) = 0 .$$

$$n x = \sum_i x_i$$

$$x = \frac{1}{n} \sum_i x_i$$

method of least squares.

population deviation

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2}$$

$$n \sigma^2 = \sum_{i=1}^{n} x_i^2 - 2 \langle x \rangle \sum_i x_i + n \langle x \rangle^2$$

$$= n \left(\langle x^2 \rangle - \langle x \rangle^2 \right)$$

$$\therefore \quad \sigma^2 = \langle x^2 \rangle - \langle x \rangle^2 .$$

$$\sigma^2 = \sum_j (x_j - \langle x \rangle)^2 p_j$$

$$= \int (x - \langle x \rangle)^2 f(x)\, dx$$

$$X \qquad Y = aX + b .$$

$$\langle Y \rangle = a \langle x \rangle + b .$$

$$\sigma^2 (Y) = a^2 \int (\; x \; - \; \langle x \rangle - \;)^2 f(x)\, dx .$$

$$= a^2 \sigma^2 (X)$$

Moments of Probability Distributions.

矩.

$$\langle (x-\langle x\rangle)^k \rangle = \sum_j (x_j - \langle x\rangle)^k p_j$$
$$= \int_{-\infty}^{+\infty} (x-\langle x\rangle)^k f(x)\,dx$$

moment-generating function.

$$\langle e^{tx}\rangle = \int \boxed{e^{tx}} f(x)\,dx = 1 + t\underline{\langle x\rangle} + \frac{t^2}{2!}\langle x^2\rangle.$$

Therefore

$$\langle x\rangle = \left.\frac{d\langle e^{tx}\rangle}{dt}\right|_{t=0}$$
$$\langle x^2\rangle = \left.\frac{d^2\langle e^{tx}\rangle}{dt^2}\right|_{t=0}.$$

example

$\boxed{1}$ $\boxed{2}$ $\boxed{3}$ $\boxed{4}$.

$P(3) = \frac{1}{6}$ $\quad P(4) = \frac{1}{6}$

$P(5) = \frac{1}{3}$ $\quad P(6) = \frac{1}{6}$ $\quad P(7) = \frac{1}{6}$

$$M = \langle e^{tx}\rangle = \frac{1}{6}\left(e^{3t} + e^{4t} + 2e^{5t} + e^{6t} + e^{7t}\right).$$
$$M' = \frac{1}{6}\left(3e^{3t} + 4e^{4t} + 10e^{5t} + 6e^{6t} + 7e^{7t}\right)$$
$$M'' = \frac{1}{6}(9 + 16 + 50 + 36 + 49)$$

$M'(t=0) = 5. = \langle x\rangle$ $\qquad \langle x^2\rangle = \frac{80}{3}$.

$$\sigma^2 = \langle x^2 \rangle - \langle x \rangle^2 = \frac{80}{3} - 5 = \frac{5}{3}.$$

Covariance and Correlation

$$f(x,y) = f(x)\, g(y) \Leftrightarrow \text{independent}.$$

$$\mathrm{cov}(X,Y) = \langle (X - \langle X \rangle)(Y - \langle Y \rangle) \rangle$$

independent.

$$\int \cancel{(X - \langle X \rangle)} f(x)\, dx \int \cancel{(Y - \langle Y \rangle)} g(y)\, dy = 0.$$

normalized covariance : correlation

$$-1 \le \frac{\mathrm{cov}(X,Y)}{\sigma(X)\,\sigma(Y)} \le 1$$

proof:

$$Q = \langle [a(X - \langle X \rangle) + c(Y - \langle Y \rangle)]^2 \rangle \ge 0.$$

$$= a^2 \underbrace{\langle (X - \langle X \rangle)^2 \rangle}_{\sigma^2(X)} + 2ac\, \mathrm{cov}(X,Y) + c^2 \underbrace{\langle (Y - \langle Y \rangle)^2 \rangle}_{\sigma^2(Y)}$$

its discriminant

$$4c^2 \mathrm{cov}^2(X,Y) - 4c^2 \sigma^2(X)\, \sigma^2(Y) \le 0.$$

$$\mathrm{cov}^2(X,Y) \le \sigma^2(X)\, \sigma^2(Y)$$

$$Y = aX + b$$

$$\frac{\mathrm{cov}(X, aX+b)}{\sigma(X)\,\cancel{a}\,\sigma(X)} = \langle (X-\langle X\rangle)(\cancel{a}X - \cancel{a}\langle X\rangle)\rangle$$

$$= \pm 1$$

$$Y = aX + b \Leftrightarrow \pm 1.$$

Marginal Probability Distribution

$$\boxed{F(x)} = \int f(x, y)\, dy$$

$$\boxed{G(y)} = \int f(x, y)\, dx$$

Conditional Probability Distribution.

$$P(X = x \mid Y = y_0)$$

$$f(x, y_0).$$

Binomial Distribution.

example. $6.$ $p = \frac{1}{6}$

4 times.

$$P(S = s) = \binom{4}{s} \underline{a}^s b^{4-s} \qquad a + b = 1.$$

$$P(S = s) = \binom{n}{s} p^s q^{n-s} \qquad p + q = 1.$$

example. moment-generating function

$$\langle e^{tS} \rangle = \langle e^{t(S_1 + S_2 \cdots + S_n)} \rangle$$
$$= \langle e^{tS_1} \rangle \langle e^{tS_2} \rangle \langle e^{tS_3} \rangle \cdots$$
$$= \left(\langle e^{tS_1} \rangle \right)^n$$

$$\langle e^{tS_1} \rangle = p e^{t} + q$$

$$\langle e^{tS} \rangle = (p e^{t} + q)^n.$$

$$\frac{\partial \langle e^{tS} \rangle}{\partial t}\Big|_{t=0} = np\underline{e^{t}}\, \underline{(pe^{t} + q)^{n-1}}\Big|_{t=0} = np.$$

$$\frac{\partial^2 \langle e^{tS} \rangle}{\partial t}\Big|_{t=0} = npe^{t}(pe^{t} + q)^{n+1} + n(n-1)p^2 e^{2t}(pe^{t} + q)^{n-1}\Big|_{t=0}$$

$$\langle S^2 \rangle = np + n(n-1)p^2.$$

$$\sigma^2(S) = \langle S^2 \rangle - \langle S \rangle^2 = npq.$$

Poisson Distribution.

example

Decay, Poisson noise.

reapted at a constant rate of probability.

$$P_n(t+dt) = P_n(t)\,P_0(dt) + P_{n-1}(t)\,P_1(dt)$$

$$P_1(dt) = \mu\, dt \ll 1 \qquad P_0(dt) = (1-\mu)\,dt$$

$$\frac{dP_n(t)}{dt} = \frac{P_n(t+dt) - P_n(t)}{dt} = \mu P_{n-1}(t) - \mu P_n(t)$$

$n=0$

$$\frac{dP_0(t)}{dt} = -\mu P_0(t) \Rightarrow P_0(t) = e^{-\mu t}$$

$$P_n(t) = \frac{(\mu t)^n}{n!} e^{-\mu t} \qquad P_n(0) = \delta_{n0}$$

$\mu t \to \mu$

$$\boxed{P(n) = \frac{\mu^n}{n!} e^{-\mu}}$$ Poisson distribution

$$\langle x \rangle = \left(\sum_{n=1}^{\infty} n \cdot \frac{\mu^n}{n!}\right) e^{-\mu} = \cancel{e^{-\mu}} \cdot \mu \sum_{n=1}^{\infty} \frac{\mu^{n-1}}{(n-1)!} = \mu$$

$$\langle x^2 \rangle = e^{-\mu} \sum \left(n^2 \frac{\mu^n}{n!}\right) = e^{-\mu} \sum_{n=1}^{\infty} \left[\frac{\cancel{\mu^n e^{\mu}}}{(n-2)!} + \frac{\mu^n}{(n-1)!} \right]$$

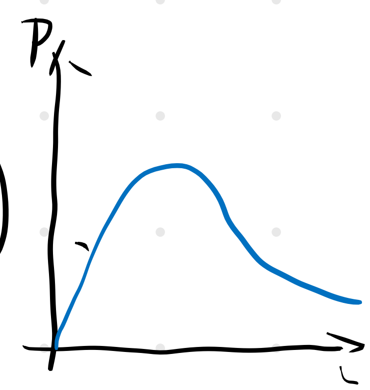
$$= \mu^2 + \mu$$

$$\sigma^2 = \mu$$

$$\langle e^{tx} \rangle = \sum_{n=0}^{\infty} \frac{\mu^n}{n!} e^{-\mu} e^{tn} = e^{-\mu} \sum_{n=0}^{\infty} \frac{(\mu \cdot e^t)^n}{n!}$$

$$= e^{-\mu} \cdot e^{\mu \cdot e^t}$$

$$= \exp(\mu(e^t - 1))$$



Relation to Binomial Distribution

$$P(S=s) = \frac{\boxed{n!}}{s!\,\boxed{(n-s)!}} p^s \boxed{q^{n-s}} \qquad \begin{array}{l} n \to +\infty \\ p \to 0 \end{array} \qquad np \to \mu$$

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

$$\frac{n!}{(n-s)!} \sim \left(\frac{n}{e}\right)^n \cdot \left(\frac{e}{n-s}\right)^{n-s} \sim \left(\frac{n}{e}\right)^s \left(\frac{n}{n-s}\right)^{n-s}$$

$$= \left(\frac{n}{e}\right)^s \boxed{\left(1+\frac{s}{n-s}\right)^{n-s}} \to e^s$$

$$= n^s$$

$$q^{n-s} = (1-p)^{n-s} \qquad p = \frac{\mu}{n}$$

$$= \left(1-\frac{\mu}{n}\right)^n \underline{\left(1-\frac{\mu}{n}\right)^{-s}} \to 1$$

$$= e^{-\mu}$$

$$P(S=s) = \frac{n!}{s!(n-s)!} p^s q^{n-s}$$

$$\sim \frac{n^s}{s!} p^s e^{-\mu}$$

$$= \boxed{\frac{\mu^s}{s!} e^{-\mu}} \quad \text{Poisson distribution}$$

Gauss' Normal Distribution

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

limits of Poisson and Binominal Distributions

$$p(n) = \frac{\mu^n}{n!} e^{-\mu}.$$

$$\ln p(n) = \underline{n \ln \mu} - \mu - \ln\sqrt{2n\pi} \; \underline{- n \ln n} + n$$

$n \to \mu + v$.

$$= (\mu + v) \ln\left(1 - \frac{v}{\mu+v}\right) + v - \ln\sqrt{2(\mu+v)\pi}.$$

$\ln(1-x)$
$\downarrow$
Σ.

$$= -\sum_{t=1}^{\infty} \frac{v^t}{t(\mu+v)^{t-1}} + v - \ln\sqrt{2\pi(\mu+v)}$$

$$\ln p(n) \approx -\frac{v^2}{2\mu} - \ln\sqrt{2\pi\mu}$$

$$p(n) \sim \frac{1}{\sqrt{2\pi\mu}} e^{-\frac{v^2}{2\mu}}.$$

Transformations of Random Variables.

$$X \to Y \qquad y = y(x).$$

$$\boxed{f(x)\,dx = g(y)\,dy} \Rightarrow g(y) = f(x) \cdot \frac{dx}{dy}.$$

$$\underset{f}{X}, \underset{g}{Y} \to \underset{u}{U(X,Y)} \quad \underset{v}{V(X,Y)}$$

$$g(u,v) = f(x(u,v), y(u,v)) |J|.$$

$$|J| = \left|\frac{\partial(x,y)}{\partial(u,v)}\right| = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}.$$

Addition

$$x = x \qquad J = 1$$
$$z = x + y .$$

$$g(z) = \int_{-\infty}^{+\infty} f_1(x) f_2(z-x)\, dx$$
$$= \sqrt{2\pi}\,(f_1 * f_2)(z).$$

$$\langle e^{itx} \rangle = \boxed{\int e^{itx} f(x)\, dx} = \sqrt{2\pi}\, f^T(t).$$

$$g^T(t) = \sqrt{2\pi}\, f_1^T(t) f_2^T(t)$$

$$g(z) = \int e^{-izt} f_1^T(t) f_2^T(t)\, dt$$

$$\langle e^{itz} \rangle = \langle e^{itx} \rangle \langle e^{itY} \rangle \qquad (z = X + Y)$$

Multiplication

$$x = x \qquad J = \frac{1}{x}.$$
$$z = xy .$$

$$g(z) = \int_{-\infty}^{+\infty} f\left(x, \frac{z}{x}\right) \frac{dx}{|x|}$$
$$= \int f(x) f_2\left(\frac{z}{x}\right) \frac{dx}{x}$$

Gamma Distribution

$$X \sim N(0, \sigma^2)$$

$$Y = X^2 \qquad f(x)dx = g(y)dy.$$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

$$g(y) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{y}{2\sigma^2}} \cdot \frac{\sqrt{y}}{2} & y > 0. \\ 0 & y \leq 0 \end{cases}$$

$$z = \frac{y}{2\sigma^2}$$

$$\int g(z)dz = \frac{1}{\sqrt{\pi}} \cdot \int_0^{+\infty} z^{-\frac{1}{2}} e^{-z} dz = \frac{\Gamma(\frac{1}{2})}{\sqrt{\pi}} = 1.$$

gamma distribution.

$$g(p, \sigma; y) = \begin{cases} 0, & y \leq 0 \\ \frac{y^{p-1} e^{-\frac{y}{2\sigma^2}}}{(2\sigma^2)^p \Gamma(p)} \end{cases}$$

$$[g(p,\sigma)]^T(t) = \frac{1}{\sqrt{2\pi}} \frac{1}{(1-2i\sigma^2 t)^p}$$

$$\langle e^{itx} \rangle = \frac{1}{(1-2i\sigma^2 t)^p}$$

Statistics

error propagation

$$x_j = \bar{x} + e_j \qquad \sum_j e_j = 0$$

$$\bar{f}(x) = \frac{1}{n}\sum f(x_j) = \frac{1}{n}\sum f(\bar{x} + e_j)$$

$$= f(\bar{x}) + \cancel{\frac{1}{n} f'(\bar{x}) \boxed{\sum e_j}} + \frac{1}{2n} f'' \sum e_j^2$$

$$= f(\bar{x}) + \frac{1}{2}\sigma^2 f''(\bar{x}) + \cdots$$

$$\sigma^2(f) = \frac{1}{n}\sum_j (f_j - \bar{f})^2 \approx \frac{1}{n}(f'(\bar{x}))^2 \sum e_j^2$$

$$= [f'(\bar{x})]^2 \sigma^2.$$

$$f(\bar{x} \pm \sigma) = f(\bar{x}) \pm f'(\bar{x})\sigma$$

two variant.

$$\sigma^2(f) = \frac{1}{rs}\sum\sum (f_{jk} - \bar{f})^2 \qquad f_x = \frac{\partial f}{\partial x}.$$

$$= \frac{1}{rs}\sum\sum (f_x u_j + f_y v_k)^2.$$

$$= \frac{f_x^2}{r}\sum u_j^2 + \frac{f_y^2}{s}\sum v_k^2$$

$$\sigma^2(f) = f_x^2 \sigma_x^2 + f_y^2 \sigma_y^2.$$

Repeated Measurement

σ. $\qquad \frac{\sigma}{n}$

$$s^2 = \frac{1}{n}\left(\sum_j v_j^2\right) \qquad e_j = x_j - \bar{x}$$

$$= \frac{1}{n}\sum_{j=1}^{n}(e_j + \alpha)^2$$

$$= \frac{1}{n}\sum_j (e_j^2 + \alpha^2)$$

$$\alpha^2 \approx \frac{s^2}{n}$$

$$s^2\left(1 - \frac{1}{n}\right) = \boxed{\frac{1}{n}\sum e_j^2}$$

$$= \frac{1}{n}\sum_j (x_j - \bar{x})$$

$$s = \sqrt{\frac{\sum_j (x_j - \bar{x})^2}{n-1}}$$

sample standard deviation

The χ^2 Distribution

$$\chi^2 := \sum_{j=1}^{n}\boxed{\left(\frac{u_j - u(t_j, a, \dots)}{\sigma_j}\right)^2}$$

a	b	a+b
c	d	c+d
a+c	b+d	n

$$u_j \sim N$$

$$\chi^2 = \sum_{j=1}^{n} X^2 \qquad \boxed{X \sim N}$$

$$g(\chi^2 = y) = \frac{y^{\frac{n}{2}-1} e^{-\frac{y}{2}}}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})}$$

$$\boxed{P(\chi^2 > y_0) \int_{y_0}^{\infty} g(y)\,dy}$$

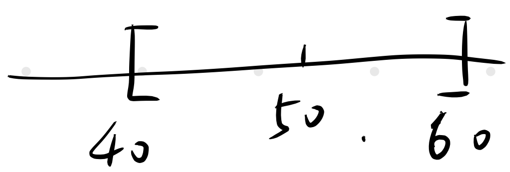
Student t Distribution. W. S. Gosset.

$$T = \frac{Y\sqrt{n}}{\sqrt{\frac{S}{n}}}$$

$$Y = \frac{1}{n}\sum_j x_j - \mu \qquad X_j \sim N.$$

$$f_n(T) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi n}\,\Gamma(\frac{n}{2})}\left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}.$$

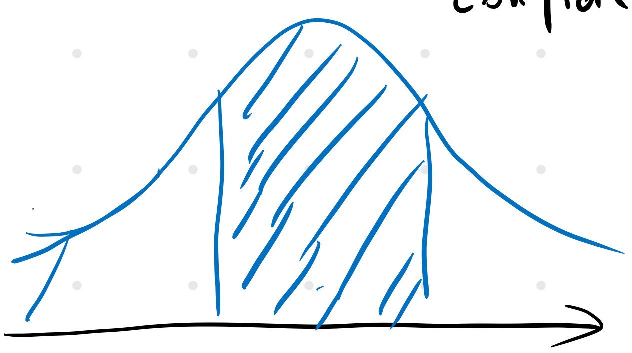
Confidence interval.



$X = 0.50 \pm 0.10$ with 90% confidence.

$$P(-c_p < t < +c_p) = p.$$

$$P(-\infty < t < +c_p) = \frac{1+p}{2} \equiv \hat{p}$$



The end