Normal Distribution

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right).$$

univariate x

what if x is a <u>vector</u>. $\underline{x} = (x_1, x_2 \ldots x_n)^T$ (column vector).

$$f_X(\underline{x}) = \frac{1}{(2\pi)^{n/2}\,|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\underline{x}-\underline{\mu}_x)^T \Sigma^{-1} (\underline{x}-\underline{\mu}_x)\right)$$

Σ - covariance matrix

$$\Sigma = E\left[(\underline{X}-\underline{\mu}_x)(\underline{X}-\underline{\mu}_x)^T\right]$$

$$= \mathrm{Cov}(X_i, X_j) \quad 1 \le i,j \le n$$

$|\Sigma|$ - determinant of Σ.

Bivariate normal

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

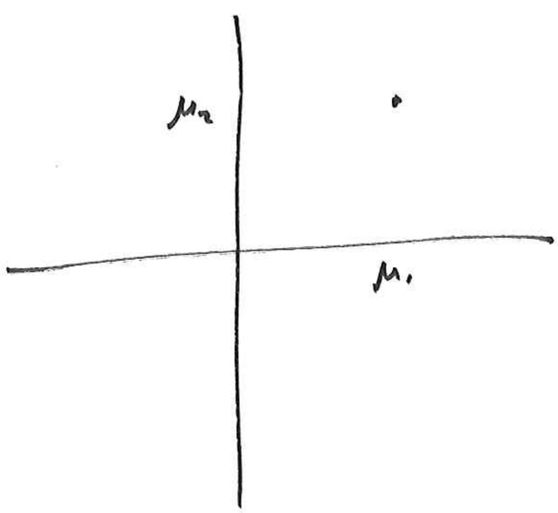
$$\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ad - bc$$

— this must be > 0 for $f_X(x)$ to define a proper distribution.

$$f(x_1, x_2) = \cancel{\frac{1}{\sqrt{2\pi}\,\sigma_x \sigma_y} -}$$

$$\frac{1}{2\pi\,\sigma_x \sigma_y \sqrt{1-\rho^2}} \cdot \exp\left(- \frac{1}{2(1-\rho^2)} \left[\frac{(x_1 - \mu_1)^2}{\sigma_1^2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2} \overset{-}{\cancel{+}} \frac{2\rho (x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1 \sigma_2} \right] \right).$$

Note should be −

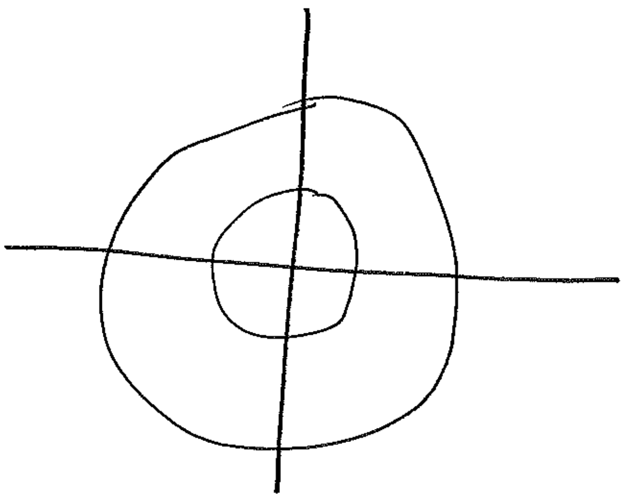


[ ] — take $\mu_1 = 0$, $\mu_2 = 0$

$$\frac{x_1^2}{\sigma_1^2} + \frac{x_2^2}{\sigma_2^2} \overset{-}{\cancel{+}} \frac{2\rho}{\sigma_1 \sigma_2} x_1 x_2$$

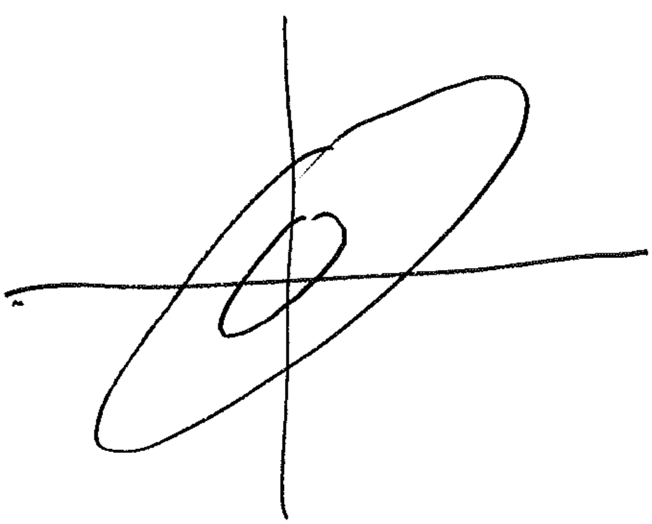
$$a x^2 + b y^2 + c x y = 1$$

– this is an ellipse.



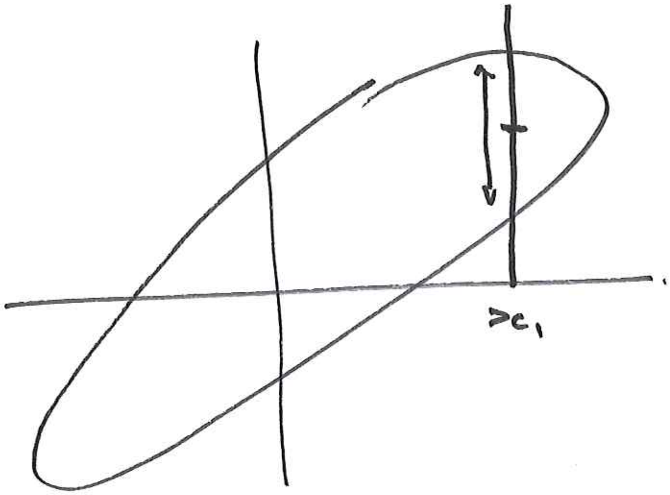
$\rho = 0$ this reduces to a product of two independent Normal distributions

contours are circles.



$\rho > 0$. contours are ellipses in the orientation shown.

conditional distribution.



$$f_{X_2|X_1}(x_2 \mid x_1)$$

Take $f(x_1, x_2)$, fix x_2 to some value.

$$\propto \exp\left(-\frac{1}{2(1-\rho^2)} \left[\cancel{\frac{x_1^2}{\sigma_1^2}} \quad \frac{x_2^2}{\sigma_2^2} - \frac{2\rho x_1 x_2}{\sigma_1 \sigma_2} \right] \right).$$

$$\exp\left(-\frac{1}{2(1-\rho^2)} \left[\frac{1}{\sigma_2^2} \left(x_2^2 - \frac{2\rho\sigma_2 x_1}{\sigma_1} x_2 \right) \right] \right).$$

$$\exp\left(-\frac{1}{2(1-\rho^2)\sigma_2^2} \left[\left(x - \frac{\rho\sigma_2 x_1}{\sigma_1} \right)^2 - \left(\frac{\rho\sigma_2 x_1}{\sigma_1} \right)^2 \right] \right)$$

$$\exp\left(-\frac{1}{2(1-\rho^2)\sigma_2^2} \left(x - \rho\frac{\sigma_2}{\sigma_1} x_1 \right)^2 \right).$$

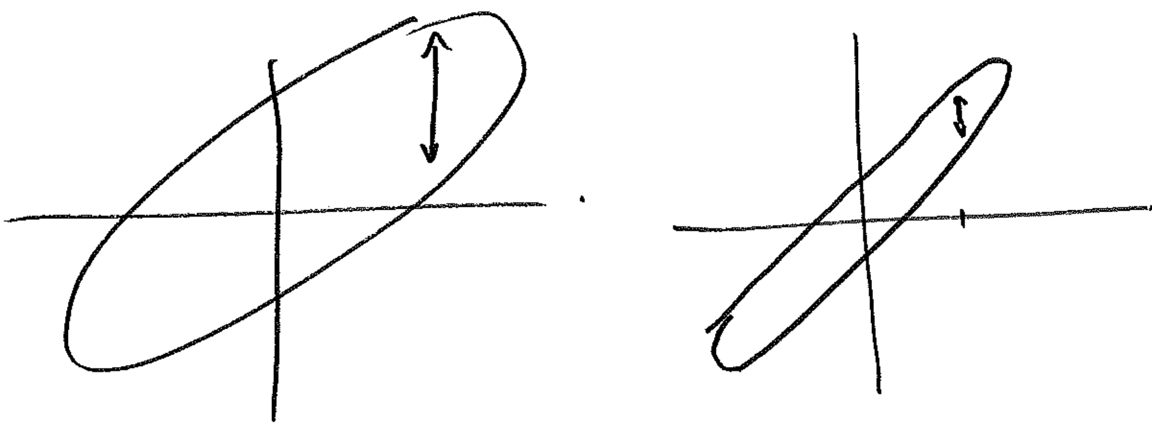
compare this with

$$\exp\left(-\frac{1}{2}\,\frac{(x-\mu)^2}{\sigma^2}\right)$$

It's a Normal, with

mean $+\ \rho\,\dfrac{\sigma_2}{\sigma_1}\,x_1$

variance $(1-\rho^2)\sigma^2$



knowing a value for x_1, the conditional distribution for $\cancel{x_2}$ $X_2 \mid X_1 = x_1$ is still Normal, but with a shifted mean, and reduced variance.

If $X_1 \ldots X_n$ are Multivariate Normal, then any linear combination $a_1X_1 + a_2X_2 + \ldots a_nX_n$ has a Normal distribution.

Law of Large Numbers / Central Limit Theorem

Let $X_1, X_2 \ldots, X_n$ be iid mean μ, variance σ^2 $\Big\{ \mu, \sigma$ exist.

Assume that otherwise the distribution of the X_i's is unknown.

Let $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$ — sample mean.

What can we say about $\bar{X}_n$ as $n \to \infty$

Law of Large Numbers

$\bar{X}_n \to \mu$ as $n \to \infty$ with probability 1

$$P\left(\lim_{n \to \infty} \bar{X}_n = \mu \right) = 1$$

As $n \to \infty$, most of the elements of the state space will have probability zero.

On the remaining element(s), $\bar{X}_n$ has the value μ.

$X_i \sim \text{Bernoulli}(p)$

$$\frac{X_1 + X_2 + \dots X_n}{n} \longrightarrow p \quad \text{with probability } 1.$$

the events where this is not true
have probability zero as $n \to \infty$.

(eg HHH... H has prob. zero as number of
tosses $\to \infty$)

Weak LLN.

for any $c > 0$

$$P(| \bar{X}_n - \mu | > c) = 0 \qquad \text{as } n \to \infty$$

if n is large enough, it is extremely likely
that $\bar{X}_n$ is very close to μ.

$\to$ Convergence in probability.

To prove this we use chebyshev's Inequality

Chebyshev's Inequality

$$P\left(|X-\mu| > a \right) \leq \frac{\mathrm{Var}(X)}{a^2} \qquad a > 0$$

to prove this we use Markov's Inequality

Markov's Inequality

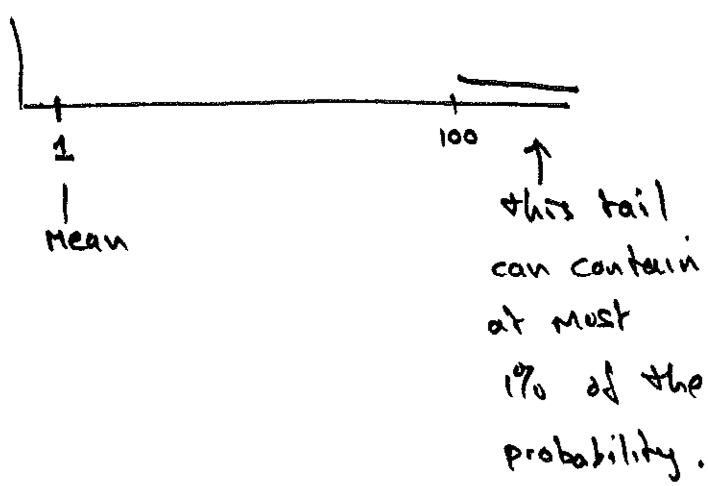
$$P\left(|X| \geq t \right) \leq \frac{E\left(|X|\right)}{t} \qquad t > 0$$

or for X taking only positive values

$$P(X \geq t) \leq \frac{E[X]}{t}$$

Example $E[X] = 1$ $\qquad$ X takes positive values.

$$P(X \geq 100) \leq \frac{1}{100}$$



In a group of 100 people
is it possible that 95% of the
people are younger than the average
age?

Yes - lots of young people and
a few old people.

Is it possible that at least 50% are
older than twice the average age?

<u>No:</u> if 50% of the people have age 2μ
then the average of <u>these</u> people is μ,
and we still have to include the other
50%

<u>Proof.</u>

for X taking +ve values.

$$P(X \geq t) \leq \frac{E[X]}{t} \qquad t > 0$$

$$E[X] = \sum_{x} x \, P(X = x)$$

$$= \sum_{x < t} x \, P(X = x) + \sum_{x \geq t} x \, P(X = x)$$

Because X takes only positive values, both of these terms are positive

$$E[X] \geq \sum_{x \geq t} x \, P(X = x)$$

$$\geq \sum_{x \geq t} t \, P(X = x)$$

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

Hence $P(X \geq t) \leq \dfrac{E[X]}{t}$

Back to Chebyshev's Inequality.

Proof.

$$P(|X-\mu| > a) = P\left((X-\mu)^2 \cancel{\geq} > a^2\right)$$

$$\leq \frac{E\left[(X-\mu)^2\right]}{a^2} \quad \text{by Markov's Inequality.}$$

$$= \frac{Var(X)}{a^2}$$

Hence

$$P(|X-\mu| > a) \leq \frac{Var(X)}{a^2}$$

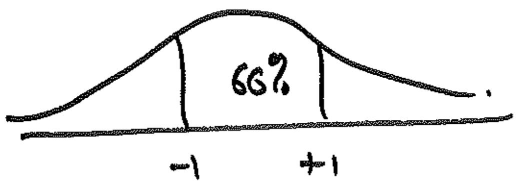
How good (close) these inequalities are depends on the distribution of X.

$$P(|X-\mu| > c\sigma)$$

$$\leq \frac{1}{c^2}$$

← prob. that RV X takes value more than c standard deviations away from the mean.

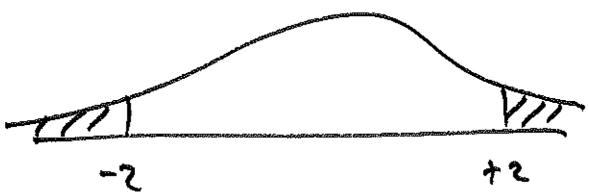
Consider Normal RV.



* prob. of X being > 1 std. dev from mean is $\simeq 0.34$

from Chebyshev's Inequality

$$P(|X-\mu| > \sigma) \leq 1.$$



for Normal dist.

prob of X being > 2 std from mean $\simeq 0.05$

from Chebyshev's inequality

$$P(|X-\mu| > 2\sigma) \leq \frac{1}{4}$$

These inequalities are most useful when we <u>don't know</u> the distribution of X.

Eg. How large must a random sample be for the probability to be at least 0.99 that the sample mean will be within 2 SD of the mean of the distribution?

$$P(|\bar{X}_n - \mu| \leq 2\sigma) \geq \frac{1}{4n}$$

Hence $1 - \frac{1}{4n} > 0.99 \Rightarrow n \geq 25$

irrespective of the distribution of X.

Back to ~~[illegible]~~ weak LLN.

$$P(|\bar{X}_n - \mu| > c) \leq \frac{Var(X)}{c^2} \quad \text{chebyshev's inequality.}$$

$$= \frac{\sigma^2}{nc^2} \longrightarrow 0 \text{ as } n \to \infty$$

what is the distribution of $\bar{X}_n$?