Lecture 24

Continuous Probability II

Recap

Continuous Probability

Continuous (uncountable) sample space. Continuous R.V.s.

$\forall x$: $\Pr[X=x]=0$. Instead, we define Prob. of intervals.

c.d.f. $F_X(x) = \Pr[X \le x]$

$$\Pr[a \le X \le b] = F_X(b) - F_X(a)$$

p.d.f. $f_X(x) = \lim_{dx \to 0} \dfrac{F_X(x+dx) - F_X(x)}{dx} = F_X'(x)$

$$\Pr[a \le X \le b] = \int_a^b f_X(x)\, dx.$$

- **Uniform dist.** in $[0,1]$ $\quad f_X(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{o.w.} \end{cases}$

- **Exponential dist** (λ) $\quad f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \ge 0 \\ 0, & \text{o.w.} \end{cases}$

  $\Pr[X > x] = e^{-\lambda x}$

Expectation of a Continuous Random Variable

If X is a r.v. with p.d.f $f_X(x)$, then the expectation of X is

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} f_X(x) \cdot x \cdot dx .$$

Justification: Use discrete approximation $Y = \frac{\lceil X \cdot n \rceil}{n}$

$$\mathbb{E}[Y] = \sum_{i=-\infty}^{\infty} \frac{i}{n} \cdot \Pr\left[\frac{i-1}{n} < X \leq \frac{i}{n}\right] \xrightarrow[n \to \infty]{} \int_{-\infty}^{\infty} x \cdot f(x) dx$$

Expectation of a Continuous Random Variable

If X is a r.v. with p.d.f $f_X(x)$, then the expectation of X is

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} f_X(x) \cdot x \cdot dx .$$

For any function $g : \mathbb{R} \to \mathbb{R}$

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) \cdot f_X(x)\, dx$$

Joint Distributions

If X and Y are continuous r.v.s then their joint p.d.f is a non-negative function $f(x,y)$

s.t. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy$

and for all $a<b$, $c<d$

$$\Pr[a \leq X \leq b, c \leq Y \leq d] = \int_c^d \int_a^b f(x,y)\,dx\,dy$$

Marginals:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy \qquad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx.$$

Joint Distributions

If X and Y are continuous r.v.s then their joint p.d.f is a non-negative function $f(x,y)$

s.t. $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy$$

and for all $a<b$, $c<d$

$$Pr\left[a \le X \le b, c \le Y \le d\right] = \int_c^d \int_a^b f(x,y)\,dx\,dy$$

Note:

$$f(x,y) = \lim_{\substack{dx \to 0 \\ dy \to 0}} \frac{Pr[x \le X \le x+dx,\ y \le Y \le y+dy]}{dx \cdot dy}$$

Independent Random Variables

X and Y are indep. continuous r.v.s if

for all $a < b$, $c < d$

$$\Pr[a \le X \le b,\ c \le Y \le d] = \Pr[a \le X \le b] \cdot \Pr[c \le Y \le d]$$

Equivalently, if $f_{XY}(x, y) = f_X(x) \cdot f_Y(y)$.

for all x, y.

Linearity of Expectation

If $X_1, \dots, X_n$ continuous r.v.s then

$$\mathbb{E}[X_1 + \dots + X_n] = \mathbb{E}[X_1] + \dots + \mathbb{E}[X_n].$$

Independent Random Variables

1. If X and Y are indep. then $\mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$.

2. If X and Y are indep. then

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y).$$

Example

$X \sim Exp(\lambda)$ $Y \sim Exp(\mu)$ X, T indep.

what's $Pr[X > Y] = ?$

Example

$X \sim Exp(\lambda) \qquad Y \sim Exp(\mu) \qquad X, T$ indep.

what's $Pr[X > Y] = ?$

$$Pr[X>Y] = \int_{-\infty}^{\infty} \left[\int_{y}^{\infty} f_{X,Y}(x,y)\, dx \right] dy$$

$$= \int_{-\infty}^{\infty} \left(\int_{y}^{\infty} f_X(x) f_Y(y)\, dx \right] dy$$

$$= \int_{-\infty}^{\infty} \mu \cdot e^{-\mu y} \left[\int_{y}^{\infty} \lambda e^{-\lambda x}\, dx \right] dy$$

$$= \int_{-\infty}^{\infty} \mu\, e^{-\mu y} \cdot e^{-\lambda y}\, dy = \frac{\mu}{\mu+\lambda} \int_{-\infty}^{\infty} (\mu+\lambda) e^{-(\mu+\lambda)y} dy$$

$$= \frac{\mu}{\mu+\lambda}.$$

Example - Meeting at a restaurant

Alice & Bob try to meet for lunch between 12 and 1

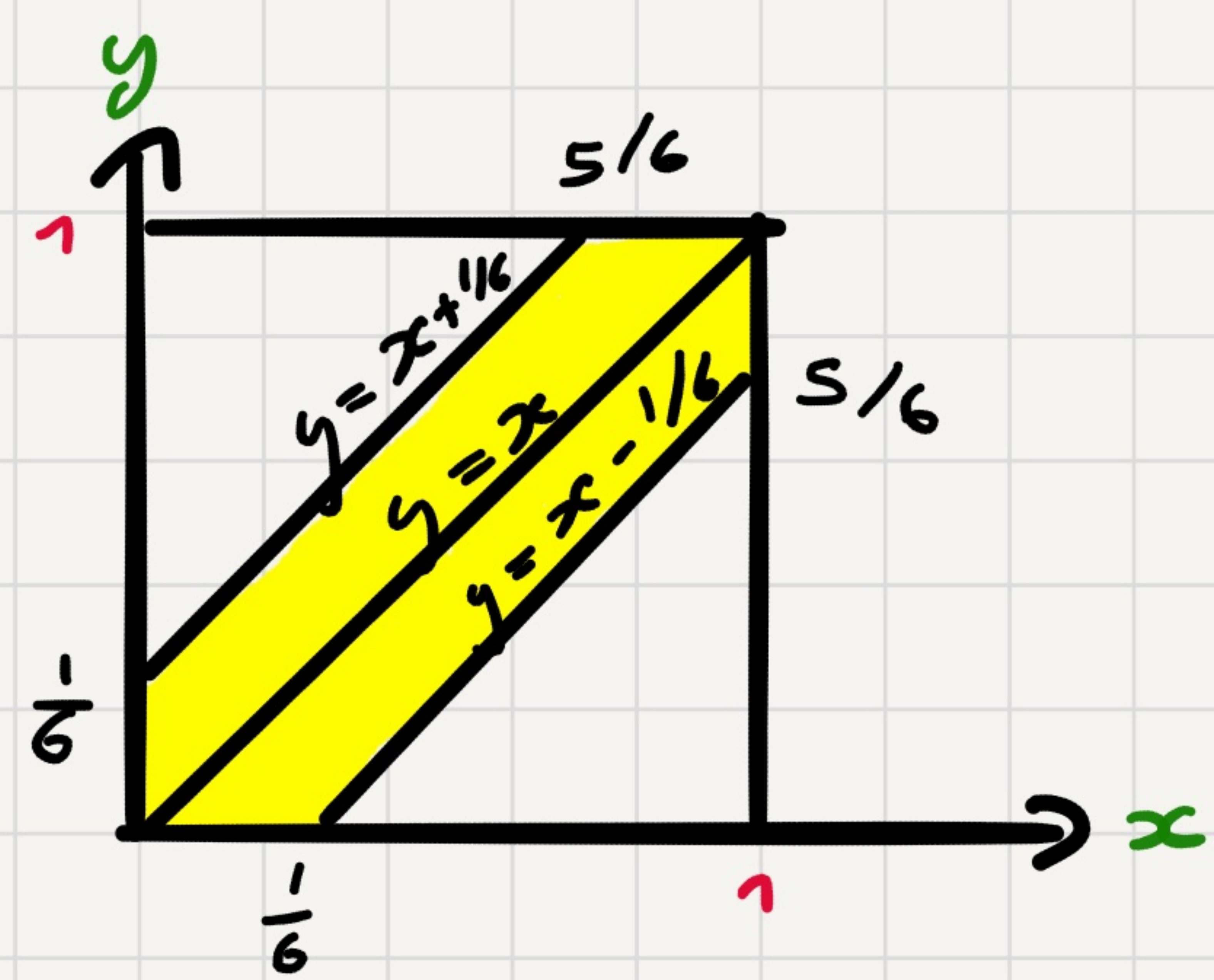
Each arrives at a uniformly random time within the hour

and will wait 10 minutes.

Q: what's the probability they meet?

X, Y indep. and $U([0,1])$

$$f_{X,Y}(x,y) = \begin{cases} 1, & \text{if } 0 \le x, y \le 1 \\ 0, & \text{o.w.} \end{cases}$$



$$\Pr[\text{meet}] = \frac{\text{area of shaded region}}{\text{area of square}} = 1 - \frac{(5/6)^2}{2} - \frac{(5/6)^2}{2} = \frac{11}{36}.$$

Normal (Gaussian) Distributions

A r.v. X is distributed $N(\mu, \sigma^2)$ if its p.d.f is

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

A r.v. X is a standard Gaussian if $X \sim N(0,1)$ and its p.d.f. is

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Normal (Gaussian) Distributions

A r.v. X is distributed $N(\mu, \sigma^2)$ if its p.d.f is

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

Theorem: If $X \sim N(0,1)$ then $Y = \sigma X + \mu$ is dist. $N(\mu, \sigma^2)$.

If $Y \sim N(\mu, \sigma^2)$ then $X = \frac{Y-\mu}{\sigma}$ is a standard Gaussian

Warmup

If X is a r.v. with pdf $f_X(x)$, $a > 0$, $b \in \mathbb{R}$

$Y = aX + b$. What's the pdf of Y?

$$F_Y(y) = \Pr[Y \leq y] = \Pr[aX + b \leq y] = \Pr\left[X \leq \frac{y-b}{a}\right]$$

$$= F_X\left(\frac{y-b}{a}\right).$$

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{d\, F_X\left(\frac{y-b}{a}\right)}{dy} = f_X\left(\frac{y-b}{a}\right) \cdot \frac{1}{a}$$

Theorem:

1. If $X \sim N(0,1)$ then $Y = \sigma X + \mu$ is dist. $N(\mu, \sigma^2)$.
2. If $Y \sim N(\mu, \sigma^2)$ then $X = \frac{Y-\mu}{\sigma}$ is a standard Gaussian

Warmup

If $Y = aX + b$

$$f_Y(y) = f_X\left(\frac{y-b}{a}\right)\frac{1}{a}$$

Proof:

1. $f_Y(y) = f_X\left(\frac{y-\mu}{\sigma}\right)\frac{1}{\sigma} = \frac{1}{\sigma}\frac{1}{\sqrt{2\pi}} e^{-(y-\mu)^2/2\sigma^2} \leftarrow N(\mu, \sigma^2)$.

2. $f_X(x) = f_Y(\sigma x + \mu) \cdot \sigma = \frac{\sigma}{\sqrt{2\pi\sigma^2}} \cdot e^{-(\sigma x + \mu - \mu)^2/2\sigma^2}$

$$= \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \leftarrow N(0,1).$$

Theorem:

1. If $X \sim N(0,1)$ then $Y = \sigma X + \mu$ is dist. $N(\mu, \sigma^2)$.
2. If $Y \sim N(\mu, \sigma^2)$ then $X = \frac{Y-\mu}{\sigma}$ is a standard Gaussian

Warmup

If $Y = aX + b$

$$f_Y(y) = f_X\left(\frac{y-b}{a}\right)\frac{1}{a}$$

Proof:

1. $f_Y(y) = f_X\left(\frac{y-\mu}{\sigma}\right)\frac{1}{\sigma} = \frac{1}{\sigma}\frac{1}{\sqrt{2\pi}} e^{-(y-\mu)^2/2\sigma^2} \leftarrow N(\mu, \sigma^2)$.

2. $f_X(x) = f_Y(\sigma x + \mu)\cdot\sigma = \frac{\sigma}{\sqrt{2\pi\sigma^2}} \cdot e^{-(\sigma x + \mu - \mu)^2/2\sigma^2}$

$$= \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \leftarrow N(0,1).$$

Expectation and Variance of Gaussian

$X \sim N(0,1)$ $\qquad f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

$$\mathbb{E}[X] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x \cdot e^{-x^2/2} \, dx$$

$$= \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{0} x \cdot e^{-x^2/2} \, dx + \int_{0}^{\infty} x \cdot e^{-x^2/2} \, dx \right) = 0$$

cancel each other.

$$\mathrm{Var}(X) = \mathbb{E}[X^2] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 \cdot e^{-x^2/2}$$

$$(-x) \cdot \left(-x \cdot e^{-x^2/2}\right)$$

$$= \frac{1}{\sqrt{2\pi}} (-x) \cdot e^{-x^2/2} \Big|_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} \, dx = 0 + 1 = 1.$$

Theorem: Let $X \sim N(\mu, \sigma^2)$ then

$$\mathbb{E}[X] = \mu$$

$$\mathrm{Var}(X) = \sigma^2$$

Proof: By previous theorems $Y = \frac{X-\mu}{\sigma}$ is standard Gaussian and has mean 0, variance 1.

$$\mathbb{E}[X] = \mathbb{E}[\sigma Y + \mu] = \sigma \mathbb{E}[Y] + \mu = 0 + \mu$$

$$\mathrm{Var}[X] = \mathrm{Var}[\sigma Y + \mu] = \sigma^2 \mathrm{Var}[Y] = \sigma^2.$$

$\blacksquare$

Two Independent Gaussians

If X, Y are independent standard Gaussians then their joint dist is $f_{XY}(x,y) = \frac{1}{2\pi} \cdot e^{-(x^2+y^2)/2}$

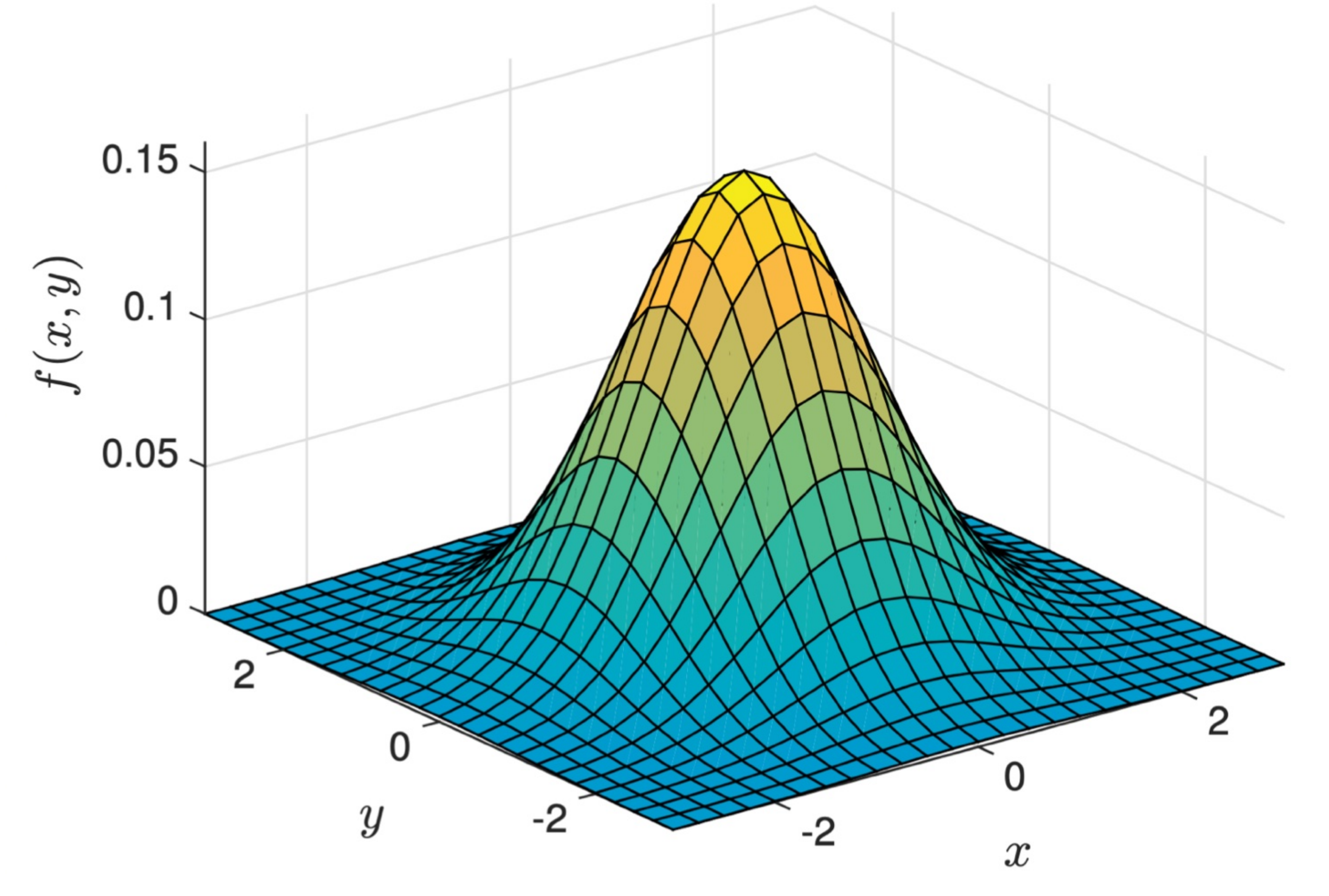


Figure 3: The joint density function $f(x,y) = \frac{1}{2\pi} e^{-(x^2+y^2)/2}$ is rotationally symmetric.

Two Independent Gaussians

If X, Y are independent standard Gaussians then their joint dist is $f_{XY}(x,y) = \frac{1}{2\pi} \cdot e^{-(x^2+y^2)/2}$

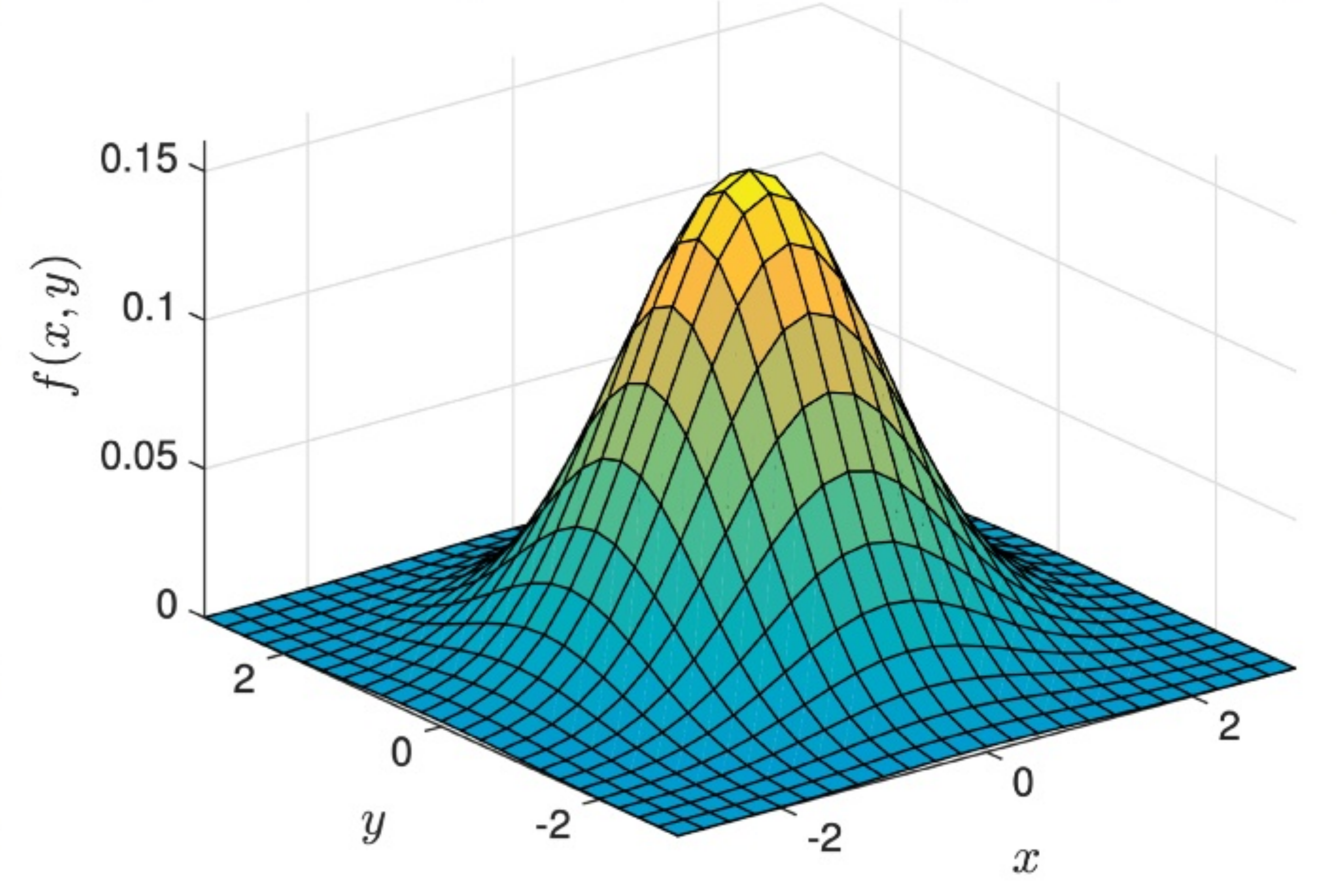
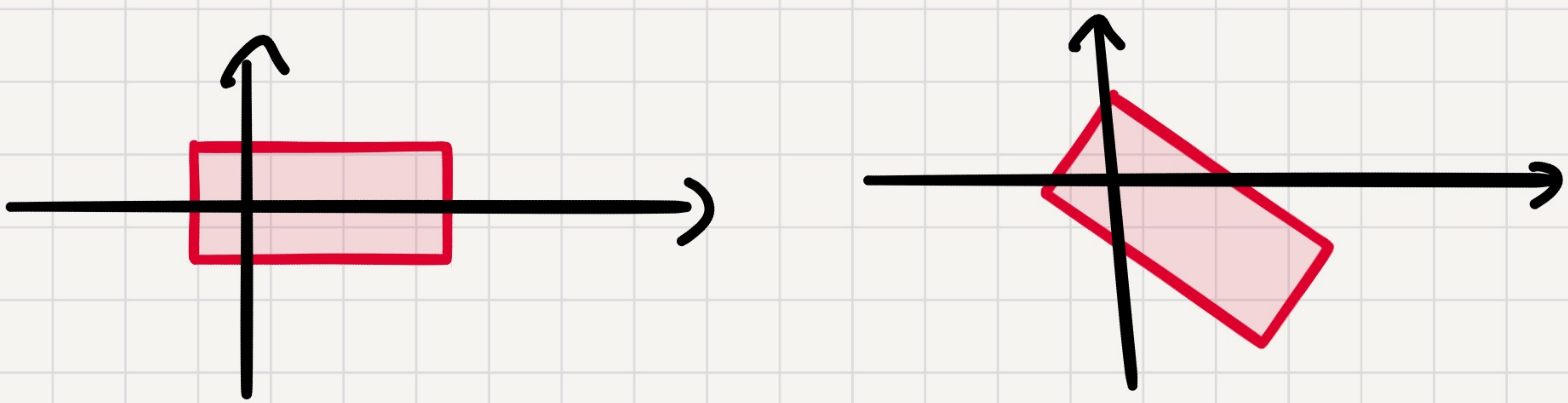


Figure 3: The joint density function $f(x,y) = \frac{1}{2\pi}e^{-(x^2+y^2)/2}$ is rotationally symmetric.

Observation: The joint distribution is rotation symmetric.

- depends only on radius, not on angle.

e.g.,



the two events occur with the same probability

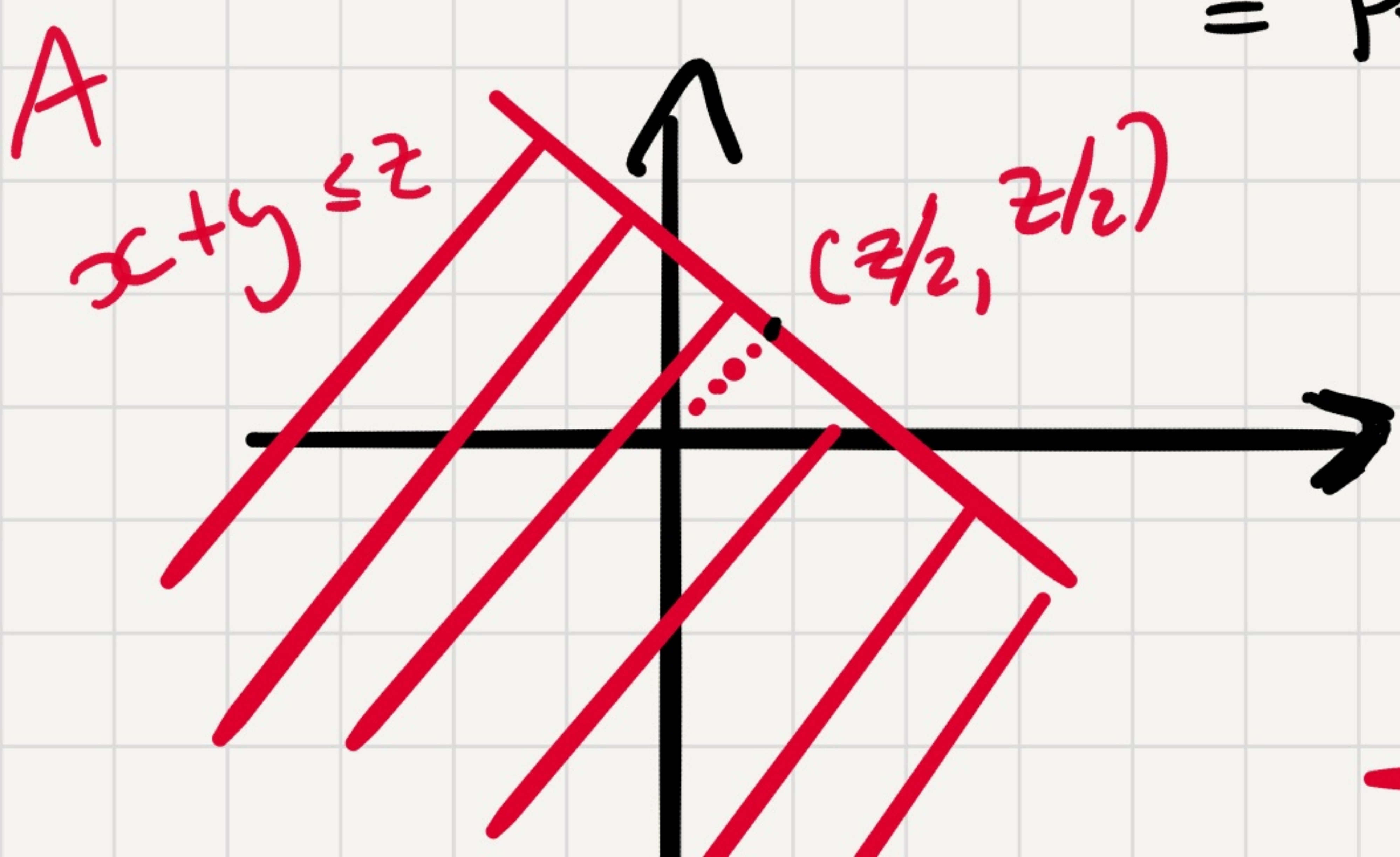
Sum of Independent Gaussians is a Gaussian

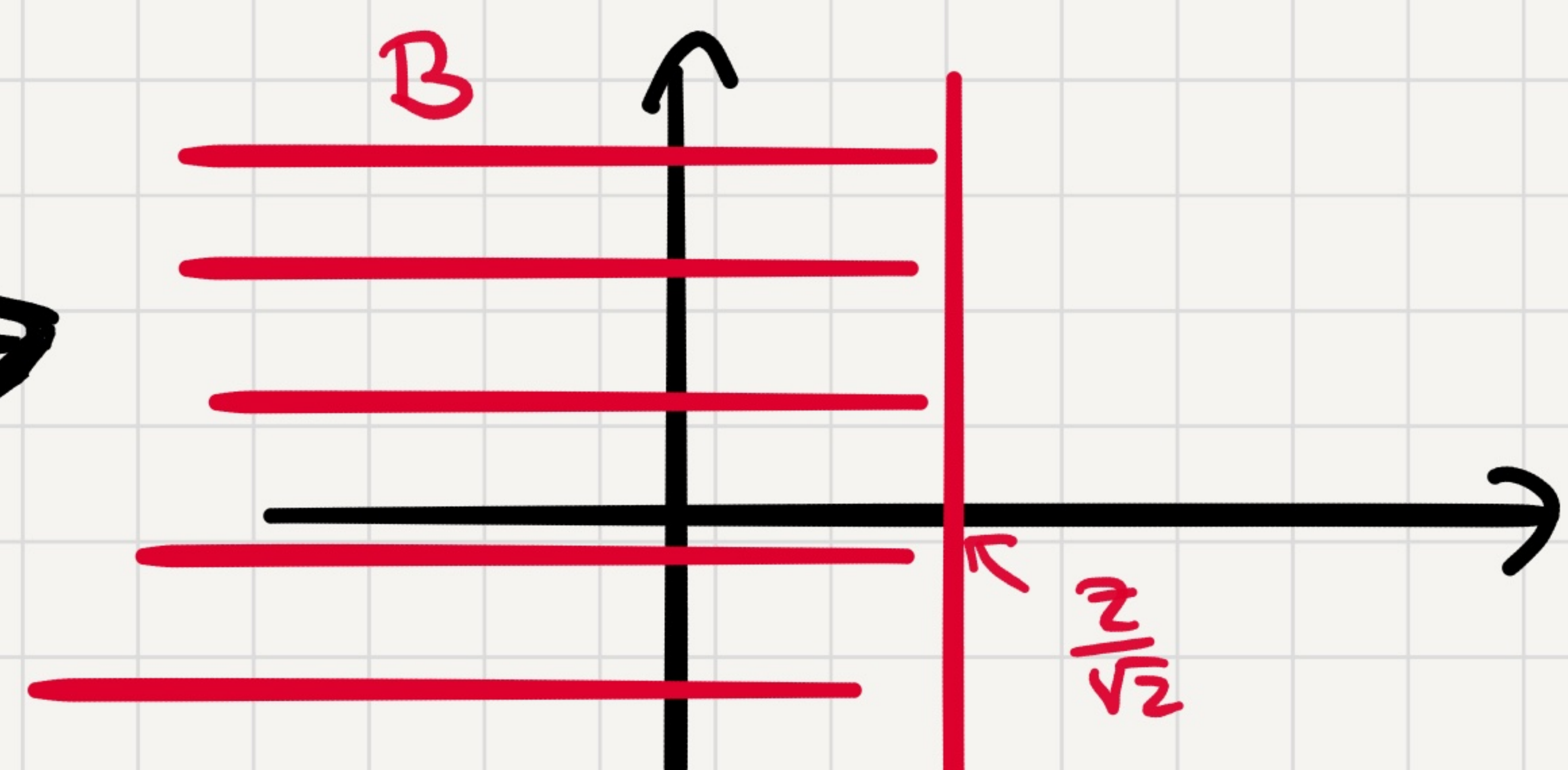
Theorem: If $X \sim N(0,1)$ and $Y \sim N(0,1)$ are independent

then $X + Y \sim N(0,2)$

Proof: Let $Z = X + Y$

$$\Pr[Z \le z] = \Pr[X + Y \le z]$$
$$= \Pr[(X,Y) \in A]$$
$$= \Pr[(X,Y) \in B] = \Pr\left[X \le \tfrac{z}{\sqrt{2}}\right]$$





Sum of Independent Gaussians is a Gaussian

Theorem: If $X \sim N(0,1)$ and $Y \sim N(0,1)$ are independent

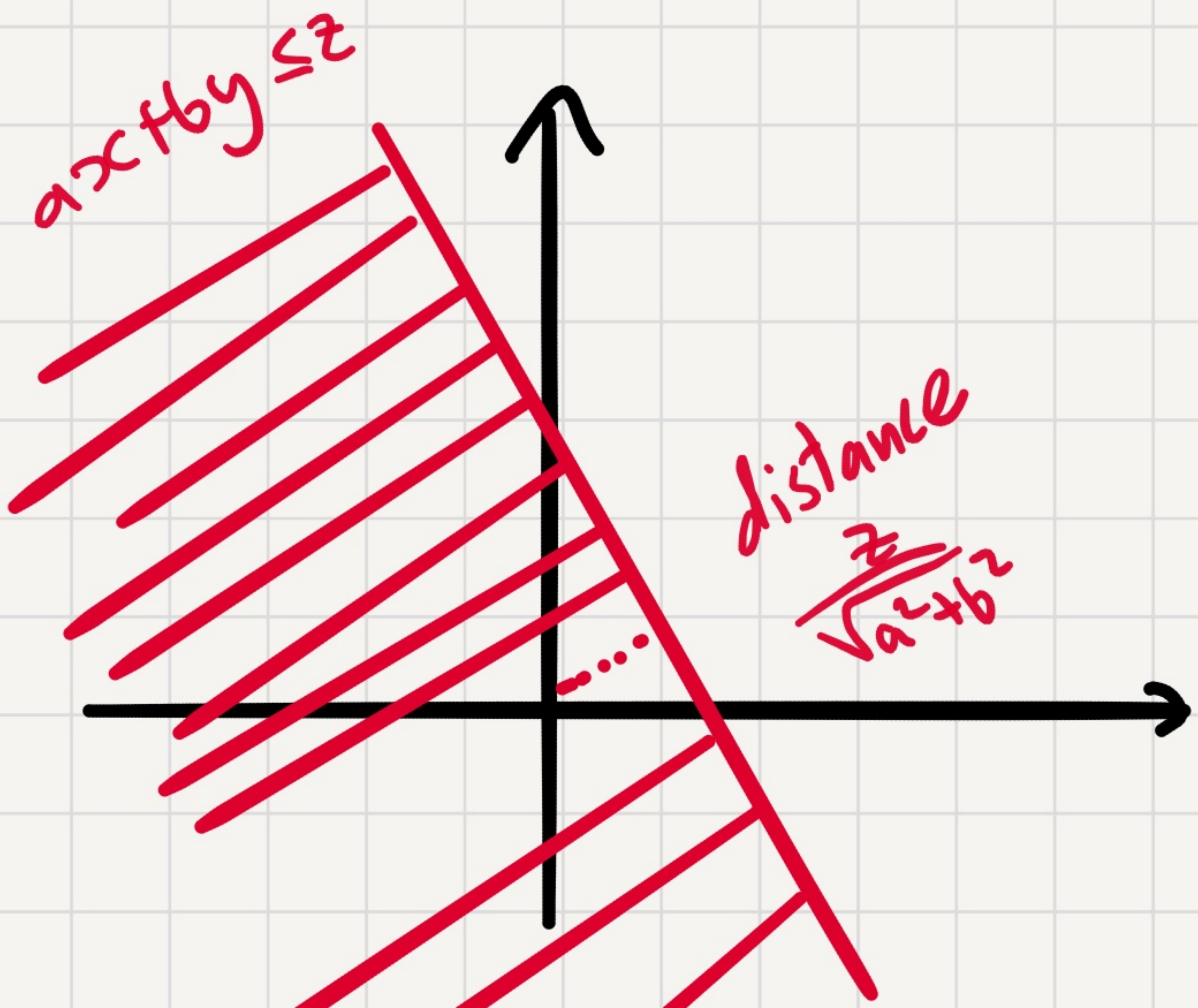
then $aX + bY \sim N(0, a^2+b^2)$

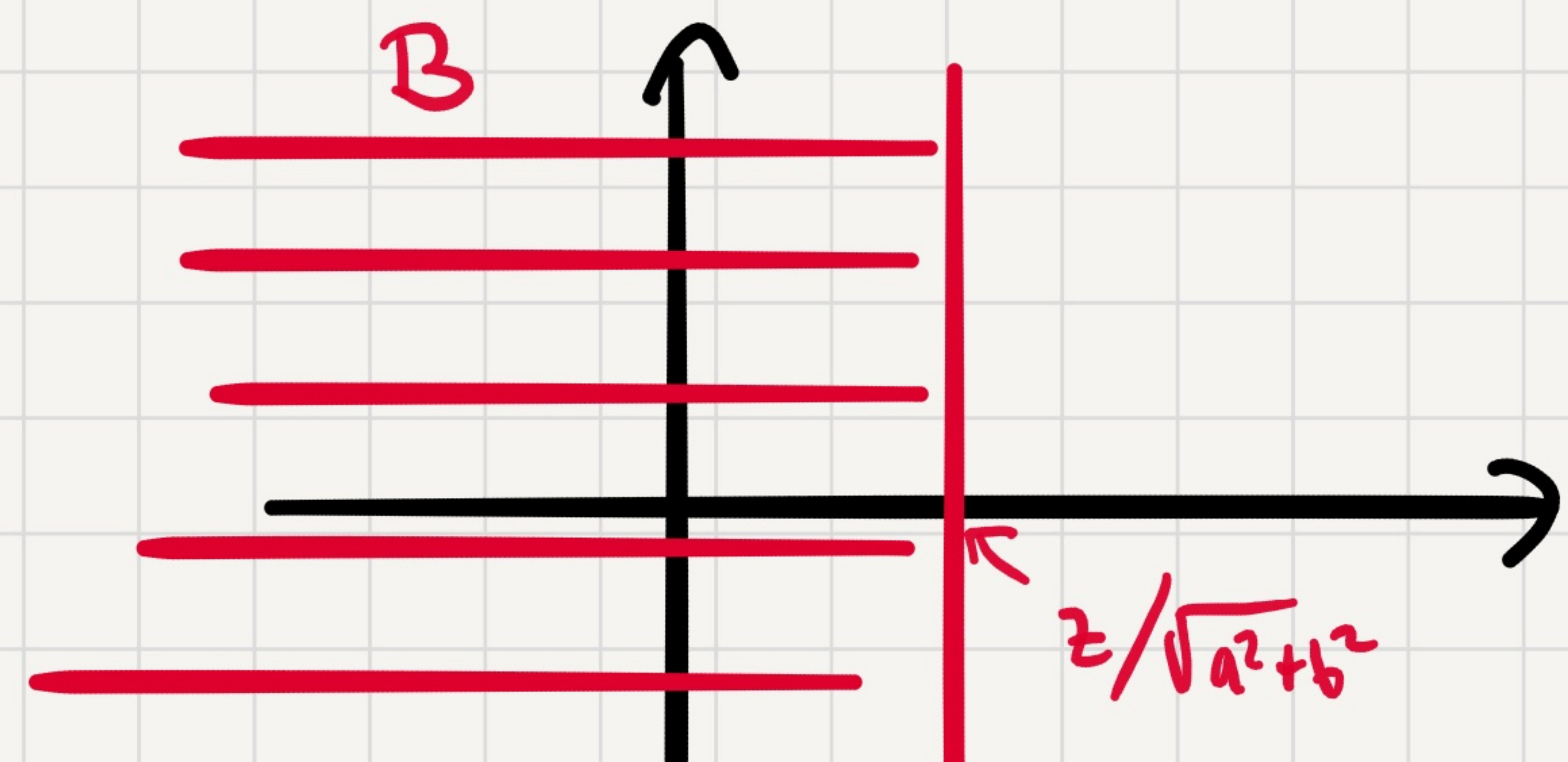
Proof: Let $Z = aX + bY$

$$\Pr[Z \le z] = \Pr[aX + bY \le z]$$

$$= \Pr[(X,Y) \in A]$$

$$= \Pr[(X,Y) \in B] = \Pr\left[X \le \frac{z}{\sqrt{a^2+b^2}}\right]$$





Sum of Independent Gaussians is a Gaussian

Theorem: If $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$ are indep. then $X+Y \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$.

Proof: $\hat{X} = \frac{X-\mu_1}{\sigma_1}$ $\quad \hat{Y} = \frac{X-\mu_2}{\sigma_2}$

$\hat{X}, \hat{Y}$ indep. standard Gaussians

$$X+Y = \sigma_1\hat{X} + \mu_1 + \sigma_2\hat{Y} + \mu_2$$

$$= \underbrace{(\mu_1+\mu_2) + \underbrace{\sigma_1\hat{X} + \sigma_2\hat{Y}}_{N(0,\,\sigma_1^2+\sigma_2^2)}}_{N(\mu_1+\mu_2,\,\sigma_1^2+\sigma_2^2).}$$

Recap

Weak Law of Large Numbers

If $X_1, X_2, \ldots$ are i.i.d.

with $\mathbb{E}X_i = \mu$ $Var(X_i) = \sigma^2$

then for all $\varepsilon > 0$

$$\Pr\left[\left| \frac{X_1 + \ldots + X_n}{n} - \mu \right| > \varepsilon \right] \xrightarrow[n \to \infty]{} 0$$

The Central Limit Theorem

If $X_1, X_2, \ldots$ are independent

with $\mathbb{E}X_i = \mu$ $\quad Var(X_i) = \sigma^2$

then $$\frac{X_1 + \cdots + X_n - \mu n}{\sigma \cdot \sqrt{n}} \xrightarrow[n\to\infty]{} N(0,1).$$

Special Case: If $X_1, X_2, \ldots$ are indep. with mean 0 variance 1

then $$\frac{X_1 + \cdots + X_n}{\sqrt{n}} \longrightarrow N(0,1)$$

The Central Limit Theorem

If $X_1, X_2, \ldots$ are independent

with $\mathbb{E}X_i = \mu$ $Var(X_i) = \sigma^2$

then $\dfrac{X_1 + \cdots + X_n - \mu n}{\sigma \cdot \sqrt{n}} \xrightarrow[n\to\infty]{} N(0,1)$.

Sanity Check:

What's the mean, var of $\dfrac{X_1 + \cdots + X_n - \mu n}{\sigma \sqrt{n}}$?

Why does it make sense that $N(0,1)$ is the limit?

$X_1, X_2, \dots,$ r.v.s with mean 0 and variance 1.

If we assume the limit $\frac{X_1 + \dots + X_n}{\sqrt{n}}$ exists, then

$$\frac{X_1 + X_3 + \dots + X_{2n-1}}{\sqrt{n}} \xrightarrow[n\to\infty]{} Z_{odd}$$

$$\frac{X_2 + X_4 + \dots + X_{2n}}{\sqrt{n}} \xrightarrow[n\to\infty]{} Z_{even}$$

$$\frac{X_1 + \dots + X_{2n}}{\sqrt{2n}} \xrightarrow[n\to\infty]{} Z$$

Z_{odd}, Z_{even}, Z have the same distribution

Z_{odd}, Z_{even} indep.

$$Z = \frac{Z_{odd} + Z_{even}}{\sqrt{2}}$$

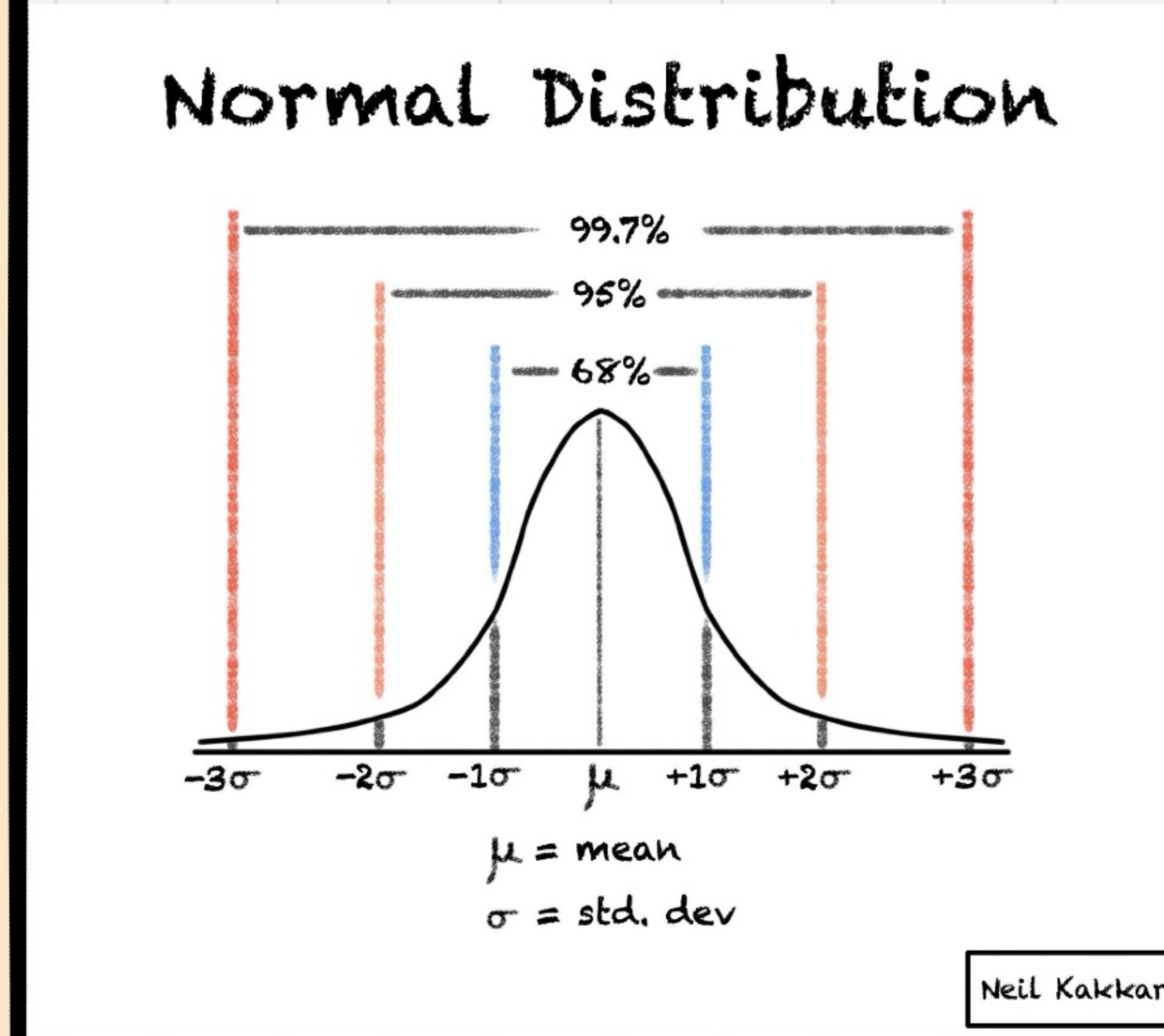
Gaussian dist. satisfies that!

The Central Limit Theorem

If $X_1, X_2, \ldots$ are i.i.d.

with $\mathbb{E}X_i = \mu$ $\quad Var(X_i) = \sigma^2$

then
$$\frac{X_1 + \cdots + X_n - \mu n}{\sigma \cdot \sqrt{n}} \xrightarrow[n\to\infty]{} N(0,1).$$



Example: $\quad X \sim Bin(n,p), \quad \mathbb{E}[X] = np \quad \sigma(X) = \sqrt{np(1-p)}$

$$P_r\left[\,|X - \mathbb{E}X| \leq k \cdot \sigma(x)\right] \xrightarrow[n\to\infty]{} \underset{Z\sim N(0,1)}{P_r\left[|Z| \leq k\right]} = \begin{cases} 0.68\ldots, & k=1 \\ 0.95\ldots, & k=2 \\ 0.997\ldots, & k=3 \end{cases}$$

More Scaling and Shifting

We saw if $Y = aX+b$

$$f_Y(y) = f_X\left(\frac{y-b}{a}\right)\cdot\frac{1}{a}.$$

Example 1: If $X \sim U([0,1])$

$Y = lX$. What's the dist. of Y?

$$d_Y(y) = \begin{cases} 1/l, & \text{if } 0 \le y \le l \\ 0, & \text{o.w.} \end{cases} \qquad U([0,l])$$

$Y = lX + b$? $\qquad U([b, b+l])$

So all uniform r.v.s are similar...

More Scaling and Shifting

We saw if $Y = aX + b$

$$f_Y(y) = f_X\left(\frac{y-b}{a}\right) \cdot \frac{1}{a}.$$

Example 2: If $X \sim \mathrm{Exp}(1)$

$Y = aX$. What's the dist. of Y?

$d_Y(y) = \frac{1}{a} \cdot e^{-y/a}$ $\qquad$ $\mathrm{Exp}(1/a)$.

$\therefore \quad Y = \frac{X}{\lambda}$ $\quad$ is $\quad \mathrm{Exp}(\lambda)$.

So all exponential r.v.s are similar...

Minimum of Two Exponentials

Let $X \sim Exp(\lambda)$, $Y \sim Exp(\mu)$ independent.

Let $Z = \min(X, Y)$. What's the dist of Z?

$$\begin{aligned} \Pr[Z \geq z] &= \Pr[\min(X,Y) \geq z] \\ &= \Pr[X \geq z, Y \geq z] \\ &= \Pr[X \geq z] \Pr[Y \geq z] \\ &= e^{-\lambda z} \cdot e^{-\mu z} = e^{-(\lambda+\mu) z} \end{aligned}$$

$Z \sim Exp(\lambda + \mu)$

Maximum of Two Exponentials

Let $X \sim Exp(\lambda)$, $Y \sim Exp(\mu)$ independent.

Let $Z = \max(X, Y)$.

Calculate $\mathbb{E}[Z] = ?$

Trick: $Z = X + Y - \min(X, Y)$.

$$\mathbb{E}[Z] = \mathbb{E}[X] + \mathbb{E}[Y] - \mathbb{E}[\min(X,Y)]$$

$$= \frac{1}{\lambda} + \frac{1}{\mu} - \frac{1}{\lambda + \mu}$$

Summary

- The Gaussian Distribution $N(\mu, \sigma^2)$

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

- $N(\mu, \sigma^2)$ is a scale & shift of $N(0,1)$.
- Joint dist. of two independent standard Gaussians is rotation symmetric.
- Sums of two independent Gaussians is a Gaussian.
- central limit theorem:

  if $X_1, X_2, \ldots$ are independent w. mean μ variance σ^2 then

$$\frac{X_1 + \cdots + X_n - \mu n}{\sigma\sqrt{n}} \xrightarrow[n\to\infty]{} N(0,1)$$