MATH 180A: INTRO TO PROBABILITY
(FOR DATA SCIENCE)

www.math.ucsd.edu/~tkemp/180A

Today: § 3.4 - 3.5

Next: § 4.1 - 4.2

Midterm 1: solutions posted; grades released. $\begin{cases} \text{Median } 76\% \\ \text{StDev } 20\% \end{cases}$

Provisional grades posted tonight.

HW3 grades released; regrade requests Tue 10/29 8am-11pm

Lab 3 grades released; regrade requests Thu 10/31 8am-11pm

HW4 now posted; due next Friday (Nov 1) by 11:59pm.

Variance

3.4

Definition: The variance of a random variable X is

$$\mathrm{Var}(X) = \mathbb{E}\left[(X-\mathbb{E}(X))^2\right].$$

I.e. first compute $\mu := \mathbb{E}(X)$. Then apply the function $g(x) = (x-\mu)^2$ to X, and compute $\mathbb{E}(g(X))$.

* If X is discrete, $\mathrm{Var}(X) = \sum_t (t-\mu)^2 \, \mathbb{P}(X=t)$

* If X is continuous, $\mathrm{Var}(X) = \int_{-\infty}^{\infty} (t-\mu)^2 \, \underline{\underline{f_X}}(t) \, dt$

In any case, $\mathrm{Var}(X) \geq 0$. Its square root is

$\sigma(X) = \sqrt{\mathrm{Var}(X)}$ is called standard deviation.

Eg. If $X \sim \mathrm{Ber}(p)$, $\mathbb{E}(X) = p$, $\therefore$

$$\mathrm{Var}(X) = \mathbb{E}((X-p)^2) = \sum_t (t-p)^2 \, \mathbb{P}(X=t)$$

$$= (1-p)^2 \, \mathbb{P}(X=1) + (0-p)^2 \, \mathbb{P}(X=0)$$

$$= (1-p)^2 p + p^2 (1-p)$$

$$= (1-p)p\,[1-p+p]$$

$$= \boxed{(1-p)p}$$

$p = \frac{1}{2}$

$\mathrm{Var} = \frac{1}{4}$

$\sigma = \frac{1}{2}$

Eg. If $U \sim \mathrm{Unif}([a,b])$, $\mathbb{E}(U) = \frac{a+b}{2}$, $\therefore$

$$\mathrm{Var}(U) = \int_{-\infty}^{\infty} \left(t - \frac{a+b}{2}\right)^2 f_U(t)\, dt = \int_a^b \left(t - \frac{a+b}{2}\right)^2 \cdot \frac{1}{b-a}\, dt$$

$f_U(t)$: $\frac{1}{b-a}$ on $[a,b]$, 0 otherwise

$$= \frac{1}{3}\left(t - \frac{a+b}{2}\right)^3 \Bigg|_{t=a}^{t=b} \cdot \frac{1}{b-a}$$

$$= \frac{1}{3} \cdot \frac{1}{b-a} \left[\left(b - \frac{a+b}{2}\right)^3 - \left(a - \frac{a+b}{2}\right)^3\right]$$

$$= \frac{1}{3} \cdot \frac{1}{b-a} \left[\left(\frac{b-a}{2}\right)^3 - \left(\frac{a-b}{2}\right)^3\right]$$

$$= \frac{1}{3 \cdot 2^3} \cdot \frac{1}{b-a} \cdot 2(b-a)^3$$

$$\frac{(b-a)^3}{12(b-a)} = \frac{(b-a)^2}{12}$$

$$\mathrm{St.\ Dev.} = \frac{|b-a|}{\sqrt{12}}$$

Variance is a measure of how "spread out from the mean" the distribution is. For example:

Theorem: Let X be a random variable with finite expectation $\mathbb{E}(X)=\mu$. Then

$$\mathrm{Var}(X)=0 \text{ iff } \mathbb{P}(X=\mu)=1.$$

(i.e. $X \equiv \mu$.)

Pf. ($\Longleftarrow$) If $\mathbb{P}(X=\mu)=1$. So X is discrete.

$\therefore \mathrm{Var}(X)=\mathbb{E}((X-\mu)^2)=(\mu-\mu)^2\,\mathbb{P}(X=\mu)=0.$

($\Longrightarrow$) [Here we will assume X is discrete, for now.]

$$0=\mathrm{Var}(X)=\underbrace{\sum_t \underbrace{(t-\mu)^2}_{\geq 0}\underbrace{\mathbb{P}(X=t)}_{\geq 0}}_{\text{sum of } \geq 0}$$

$\Rightarrow$ for each t, $(t-\mu)^2\cdot\mathbb{P}(X=t)=0.$

$\therefore$ For each t, either $\underline{(t-\mu)^2=0}$ or $\mathbb{P}(X=t)=0$.

If $t\neq\mu \Longrightarrow \mathbb{P}(X=t)=0.$

///

Chebyshev's Inequality

If X has finite mean $\mathbb{E}(X)=\mu$ and finite variance $Var(X)=\sigma^2$, then

$$\mathbb{P}(|X-\mu|\geq k\sigma) \leq \frac{1}{k^2}$$

Two ideas in the proof:

(1) Indicators. $\mathbb{1}_A = \begin{cases} 1 & \text{if } A \text{ happens} \\ 0 & \text{if } A \text{ does not} \end{cases}$

$\mathbb{1}_A$ is a 0-1 valued r.v. It is a $Ber(p)$ with $p=\mathbb{P}(A)$

$$\therefore \mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$$

(2) Monotonicity. FACT: if $X \leq Y$ then $\mathbb{E}(X) \leq \mathbb{E}(Y)$

(Intuitive; for proof, we must wait until week 8.)

Proof of Chebyshev:

$$\mathbb{P}(|X-\mu|\geq k\sigma) = \mathbb{E}\left(\mathbb{1}_{\{|X-\mu|\geq k\sigma\}}\right)$$

$$= \mathbb{E}\left(\mathbb{1}_{\left(\frac{X-\mu}{k\sigma}\right)^2\geq 1}\right)$$

$$\leq \mathbb{E}\left[\left(\frac{X-\mu}{k\sigma}\right)^2\right] = \frac{1}{k^2\cancel{\sigma^2}}\underbrace{\mathbb{E}\left[(X-\mu)^2\right]}_{\cancel{\sigma^2}}$$

$|X-\mu|\geq k\sigma$

$\Leftrightarrow \underbrace{\left(\frac{X-\mu}{k\sigma}\right)^2}_{Y}\geq 1$

if $Y\geq 1$ then $Y\geq \mathbb{1}_Y$

Normal (Gaussian) Distribution

3.5

The standard normal distribution $\mathcal{N}(0,1)$ is given by the density

$$f(t) = \underbrace{\frac{1}{\sqrt{2\pi}}\, e^{-t^2/2}}_{>0}$$

Check:

$$I = \int_{-\infty}^{\infty} e^{-t^2/2}\, dt = \sqrt{2\pi}$$

$$I^2 = \int_{-\infty}^{\infty} e^{-x^2/2}\,dx \int_{-\infty}^{\infty} e^{-y^2/2}\,dy$$

$$= \iint_{\mathbb{R}^2} \underbrace{e^{-x^2/2} e^{-y^2/2}}_{e^{-(x^2+y^2)/2}}\, dx\,dy$$

$$= \int_0^{2\pi} d\theta \int_0^{\infty} \underbrace{r\,dr\, e^{-r^2/2}}$$

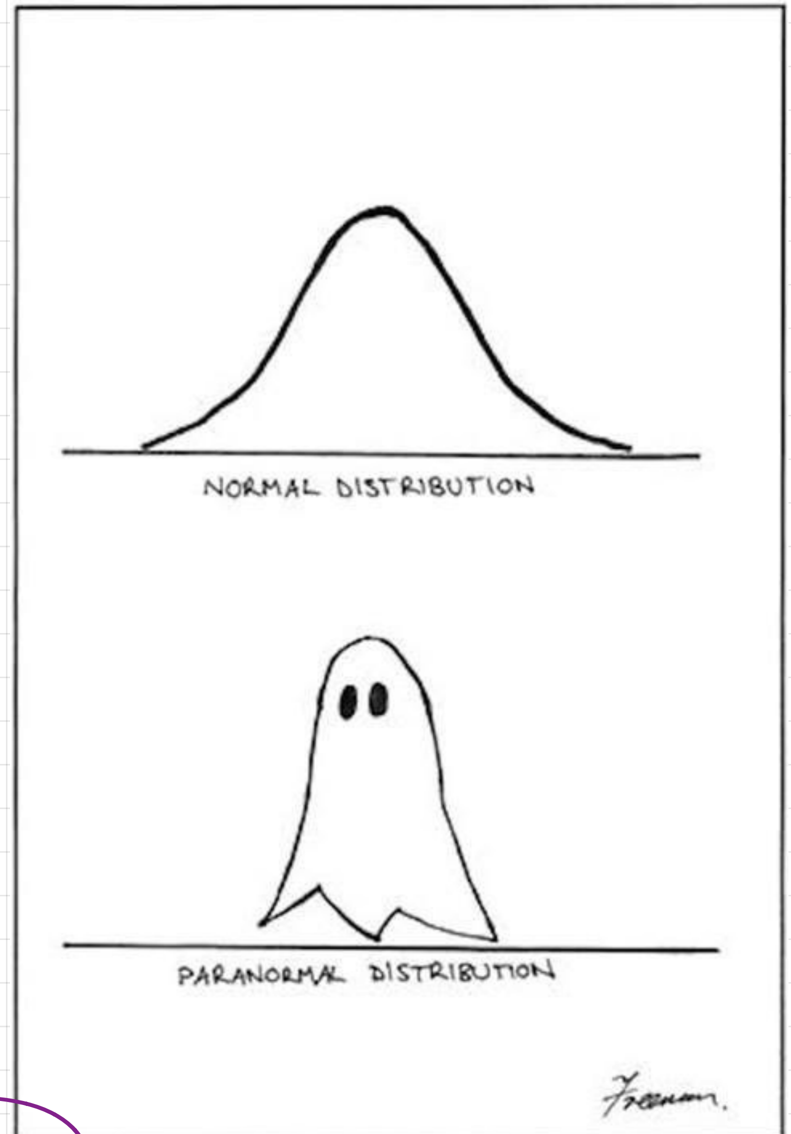
$$= 2\pi \cdot \left(-e^{-r^2/2}\right)\Big|_0^{\infty} \quad = 2\pi \cdot 1$$

Polar Coords:

$(x,y) \to (r,\theta)$

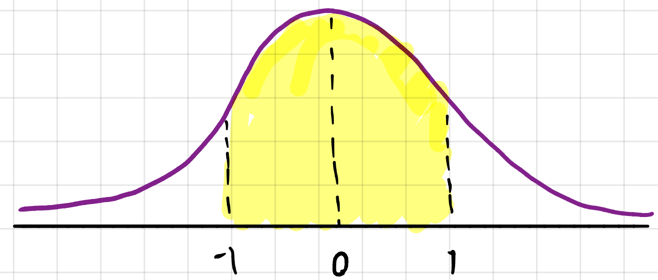
$dx\,dy \to r\,dr\,d\theta$

$x^2+y^2 = r^2$



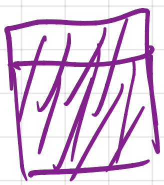
CDF of $\mathcal{N}(0,1)$

Suppose $X \sim \mathcal{N}(0,1)$. What is $\mathbb{P}(|X| \le 1)$?



$$\mathbb{P}(-1 \le X \le 1) = \int_{-1}^{1} f_X(t)\,dt = \frac{1}{\sqrt{2\pi}}\int_{-1}^{1} e^{-t^2/2}\,dt$$

$$\int_{-1}^{1}\int_{-1}^{1} \cdots$$

 ← not good for polar coords.

$$\Phi(x) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}\,dt$$