

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

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

Today: § 8.3-8.4, 9.2

Next: § 9.3

Homework 7 Due **TONIGHT!**

Lab 7 Due Wednesday, Dec 4.

Homework 8 Due Friday, Dec 6.

Final Exam: Monday, Dec 9, 11:30a-2:30p
in **REC GYM**

Flashback to Math 20C

Given two vectors $\underline{v}, \underline{w}$ in \mathbb{R}^n , their dot product is

$$\underline{v} \cdot \underline{w} = \sum_{j=1}^n v_j w_j \quad \underline{v} \cdot \underline{v} = \sum_{j=1}^n v_j^2$$

It is a positive bilinear form:

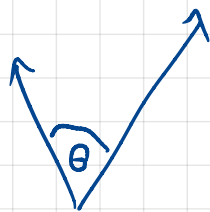
- (1) $(a\underline{u} + b\underline{v}) \cdot \underline{w} = a\underline{u} \cdot \underline{w} + b\underline{v} \cdot \underline{w}$ ($\underline{v} \cdot \underline{w} = \underline{w} \cdot \underline{v}$)
- (2) $\underline{v} \cdot (a\underline{u} + b\underline{w}) = a\underline{v} \cdot \underline{u} + b\underline{v} \cdot \underline{w}$
- (3) $\underline{v} \cdot \underline{v} > 0$ unless $\underline{v} = \underline{0}$, in which case $\underline{v} \cdot \underline{v} = 0$.

The length² of a vector is $\|\underline{v}\|^2 = \underline{v} \cdot \underline{v}$, length = $\|\underline{v}\| = \sqrt{\underline{v} \cdot \underline{v}}$

Two vectors are orthogonal if $\underline{v} \cdot \underline{w} = 0$

Cauchy-Schwarz Inequality: $|\underline{v} \cdot \underline{w}| \leq \|\underline{v}\| \|\underline{w}\|$

$$\frac{\underline{v} \cdot \underline{w}}{\|\underline{v}\| \|\underline{w}\|} = \cos \theta$$



$$\begin{aligned} -1 &\leq \frac{\underline{v} \cdot \underline{w}}{\|\underline{v}\| \|\underline{w}\|} \leq 1 \\ -1 &= \frac{\underline{v} \cdot \underline{w}}{\|\underline{v}\| \|\underline{w}\|} = 1 \text{ iff } \underline{v} = a\underline{w} \\ \text{iff } \underline{v} &= a\underline{w} \text{ for } a < 0 & \text{for some } a > 0 \end{aligned}$$

The Geometry of Random Variables

$$\frac{\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}{1}$$

Dot product \rightsquigarrow Covariance $\text{Cov}(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y)))$

\Downarrow

(Almost) positive bilinear form

- (1) $\text{Cov}(aX_1 + bX_2, Y) = a\text{Cov}(X_1, Y) + b\text{Cov}(X_2, Y)$
- (2) $\text{Cov}(X, aY_1 + bY_2) = a\text{Cov}(X, Y_1) + b\text{Cov}(X, Y_2)$
- (3) $\text{Cov}(X, X) = \text{Var}(X) > 0$ unless $X \equiv \text{const.}$
($\text{Var}(X) = 0$)

Orthogonal $\rightsquigarrow \text{Cov}(X, Y) = 0$ uncorrelated (\Leftrightarrow independent)

Length² $\rightsquigarrow \text{Var}(X)$ Length $\rightsquigarrow \sqrt{\text{Var}(X)} = \text{s.d.}(X)$

"Angles" $\rightsquigarrow \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} =: \text{Corr}(X, Y)$

Theorem: $-1 \leq \text{Corr}(X, Y) \leq 1$

$\text{Corr}(X, Y) = 0$ iff (X, Y) are uncorrelated \nRightarrow independence

$\text{Corr}(X, Y) = 1$ iff $Y = aX + b$ for some $a > 0, b \in \mathbb{R}$

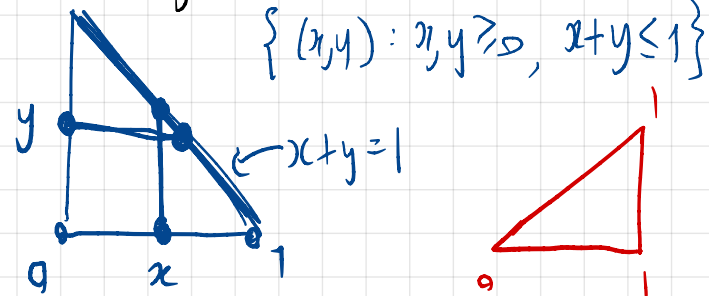
$\text{Corr}(X, Y) = -1$ iff $Y = aX + b$ for some $a < 0, b \in \mathbb{R}$.

Ex. (X, Y) uniformly distributed on the triangle  Area = $\frac{1}{2}$

$$f_{(X,Y)}(x,y) = 2 \mathbb{1}_{\Delta}(x,y)$$

$$f_X(x) = \int_0^{1-x} f_{(X,Y)}(x,y) dy$$

$$= \int_0^{1-x} 2 dy = 2(1-x) \mathbb{1}_{[0,1]}(x)$$



$X \sim Y$

$$f_Y(y) = \int_0^{1-y} 2 dx = 2(1-y) \mathbb{1}_{[0,1]}(y)$$

$$\mathbb{E}(X) = \mathbb{E}(Y) = \int_0^1 x \cdot 2(1-x) dx = \frac{1}{3} \quad \left. \begin{array}{l} \\ \end{array} \right\} \text{Var}(X) = \text{Var}(Y)$$

$$\mathbb{E}(X^2) = \mathbb{E}(Y^2) = \int_0^1 x^2 \cdot 2(1-x) dx = \frac{1}{6} \quad \left. \begin{array}{l} \\ \end{array} \right\} = \frac{1}{6} - \left(\frac{1}{3}\right)^2 = \frac{1}{18}$$

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \iint_{\Delta} xy f_{(X,Y)}(x,y) dx dy - \left(\frac{1}{3}\right)\left(\frac{1}{3}\right)$$

$$= \int_0^1 dx \int_0^{1-x} dy \, xy = \frac{1}{12} - \frac{1}{9} = -\frac{1}{36}$$

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} = \frac{-\frac{1}{36}}{\sqrt{\frac{1}{18}}\sqrt{\frac{1}{18}}} = \frac{-\frac{1}{36}}{\frac{1}{18}} = \left(-\frac{1}{2}\right)$$

Moment Generating Function Revisited

8.3

Suppose X, Y are independent, and $M_X, M_Y < \infty$ on an interval containing 0. Then

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

\uparrow fn of X \uparrow fn of Y

* If $M_Z(t) < \infty$ for all $t \in (-\delta, \delta)$, then M_Z determines the dist. of Z .
 $\Rightarrow M_{X+Y}$ determines the dist. of $X+Y$.

E.g. $X \sim \text{Poisson}(\lambda)$, $Y \sim \text{Poisson}(\mu)$ independent

$$M_X(t) = e^{\lambda(e^t - 1)} \quad M_Y(t) = e^{\mu(e^t - 1)}$$

$$M_{X+Y}(t) = e^{\lambda(e^t - 1)} e^{\mu(e^t - 1)} = e^{\lambda(e^t - 1) + \mu(e^t - 1)} = e^{(\lambda + \mu)(e^t - 1)}$$

$\Rightarrow X + Y \sim \text{Poisson}(\lambda + \mu)$

\uparrow
the MGF of $\text{Poisson}(\lambda + \mu)$

Eg. $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$ independent

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

$$\begin{aligned} X &\sim \sigma_1 Z + \mu_1, \quad \therefore M_X(t) = M_{\sigma_1 Z + \mu_1}(t) = \mathbb{E}(e^{t(\sigma_1 Z + \mu_1)}) \\ &= \mathbb{E}(e^{(\sigma_1 t)Z} e^{\mu_1 t}) \\ &= e^{\mu_1 t} \underbrace{\mathbb{E}(e^{(\sigma_1 t)Z})}_{M_Z(\sigma_1 t)} = e^{\mu_1 t} e^{\frac{(\sigma_1 t)^2}{2}} \\ &= e^{\frac{\sigma_1^2}{2} t^2 + \mu_1 t} \end{aligned}$$

$$\begin{aligned} \therefore M_{X+Y}(t) &= e^{\frac{\sigma_1^2}{2} t^2 + \mu_1 t} \cdot e^{\frac{\sigma_2^2}{2} t^2 + \mu_2 t} \\ &= e^{\frac{1}{2}(\sigma_1^2 + \sigma_2^2)t^2 + (\mu_1 + \mu_2)t} \quad \leftarrow \text{Normal} \end{aligned}$$

\approx MGF of $\mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$