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

$$\underline{v} \cdot \underline{w} =$$

It is a positive bilinear form:

(1) $(a\underline{u} + b\underline{v}) \cdot \underline{w}$

(2) $\underline{v} \cdot (a\underline{u} + b\underline{w})$

(3) $\underline{v} \cdot \underline{v}$

The length^2 of a vector is $\|\underline{v}\|^2$

Two vectors are orthogonal if

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

The Geometry of Random Variables

Dot product $\rightsquigarrow$ Covariance $\quad \mathrm{Cov}(X,Y) = \mathbb{E}\big((X-\mathbb{E}(X))(Y-\mathbb{E}(Y))\big).$

$\Downarrow$

(Almost) positive bilinear form

(1) $\mathrm{Cov}(aX_1+bX_2, Y)$

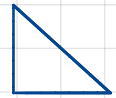
(2) $\mathrm{Cov}(X, aY_1+bY_2)$

(3) $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$

Orthogonal $\rightsquigarrow$ $\mathrm{Cov}(X,Y)=0$ uncorrelated

Length^2 $\rightsquigarrow$ $\mathrm{Var}(X)$

"Angles" $\rightsquigarrow$

Eg. (X,Y) uniformly distributed on the triangle 

$\{(x,y): x,y \geq 0,\ x+y \leq 1\}$

Moment Generating Function Revisited

8.3

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

$$M_{X+Y}(t)$$

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

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

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

9.1-9.2

Reminder: Chebyshev's Inequality

For any random variable X with finite

$$\mathbb{E}(X) = \mu \qquad Var(X) = \sigma^2$$

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

We proved this using the fact that $\mathbb{E}$ is monotone:

$$X \leq Y \Rightarrow \mathbb{E}(X) \leq \mathbb{E}(Y)$$

Eg. Ramen Menya Ultra has, on average, 1000 customers/day, with a standard deviation of 15. Estimate the probability that today they will have between 956 and 1044 customers.

(Weak) Law of Large Numbers

Let $X_1, X_2, X_3, \dots, X_n, \dots$ be an infinite sequence of i.i.d. random variables, each with $\mathbb{E}(X_j) = \mu$ and $Var(X_j) = \sigma^2$ finite.

Let $S_n = X_1 + X_2 + \dots + X_n$. Then for any fixed $\varepsilon > 0$, E.g. The Large Hadron Collider was built to detect and measure the mass of the Higgs Boson. Call the mass M.

For theoretical reasons, it is known that $M \leq 1.78 \times 10^{-23} g$.

How many trials do the LHC physicists need to do to estimate the correct mass (via sample mean) within $10^{-24} g$, with probability $\geq 95\%$?

Strong Law of Large Numbers

Let $X_1, X_2, X_3, \ldots, X_n, \ldots$ be an infinite sequence of i.i.d. random variables each with $\mathbb{E}(X_j) = \mu$.

Let $\bar{X}_n = \dfrac{X_1 + X_2 + \cdots + X_n}{n}$. Then