Math 341: Probability
Eleventh Lecture (10/20/09)

Steven J Miller
Williams College

Steven.J.Miller@williams.edu
http://www.williams.edu/go/math/sjmiller/public_html/341/

Bronfman Science Center
Williams College, October 20, 2009

Summary for the Day

Summary for the day

- Integrals related to Normal Distribution:
  - ◇ Normalization constant.
  - ◇ Gamma function and $(-1/2)!$.
- Joint density and Independence:
  - ◇ Key Lemma.
  - ◇ Example.
- Cauchy-Schwartz Inequality:
  - ◇ Proof.
  - ◇ Application to covariance.
- Clicker questions:
  - ◇ Generalizing collecting toys.
  - ◇ Three hats.

Normal Distribution and the Gamma Function

Normal distribution

Normalization constant

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} dx = 1.$$

Proof: Square the integral and use polar coordinates.

Gamma Function

$$\Gamma(s) = \int_0^\infty e^{-x} x^{s-1} dx.$$

- Useful facts:
  - Functional equation: $\Gamma(s+1) = s\Gamma(s)$.
  - Generalization of factorial: $\Gamma(n+1) = n!$.
  - Useful special value: $\Gamma(1/2) = \sqrt{\pi}$.
  - Stirling's formula:

$$\Gamma(n+1) = n! \approx n^n e^{-n} \sqrt{2\pi n} \left(1 + \frac{1}{12n} + \cdots\right),$$

Joint Density and Independence

Review

Probability Mass Function

The Probability Mass Function of a discrete random variable X is a function $f : \mathbb{R} \to [0, 1]$ given by $f(x) = \mathbb{P}(X = x)$.

Probability Density Function

The Probability Density Function of a continuous random variable X is the f such that $F(x) = \int_{-\infty}^{x} f(u)du$.

Key Lemma

Main Lemma on Independence

Random variables X and Y are independent if and only if

$$f_{X,Y}(x,y) \;=\; f_X(x) \cdot f_Y(y)$$

(i.e., their joint probability density factors).

Key Lemma

Main Lemma on Independence

Random variables X and Y are independent if and only if

$$f_{X,Y}(x,y) \ = \ f_X(x) \cdot f_Y(y)$$

(i.e., their joint probability density factors).

Proof:

- Consider box $[a,b] \times [c,d]$ with $A = \{X \in [a,b]\}$, $B = \{Y \in [c,d]\}$.
- Compare $\mathbb{P}(A)\mathbb{P}(B)$ and $\mathbb{P}(A \cap B)$.

Example

- Choose

$$f_{X,Y} = \begin{cases} (e-2)^{-1} x e^{xy} & \text{if } 0 \le x, y \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

- Let $A = \{X \le 1/2\}$ and $B = \{Y \le 1/2\}$.
- Calculation: for $0 \le x, y \le 1$:

$$f_X(x) = \frac{e^x - 1}{e-2}, \quad f_Y(y) = \frac{1 + (y-1)e^y}{(e-2)y^2}.$$

-

$$\mathbb{P}(A \cap B) = \frac{5 - 4e^{1/4}}{e-2} \approx .0947$$

$$\mathbb{P}(A)\mathbb{P}(B) = \frac{3 - 2\sqrt{e}}{4 - 2e} \cdot \frac{2\sqrt{e} - 3}{e-2} \approx .0857.$$

Cauchy-Schwartz Inequality

Standard Formulation

Cauchy-Schwartz Inequality

$$\left(\int_{-\infty}^{\infty} |f(x)g(x)|dx\right)^2 \leq \left(\int_{-\infty}^{\infty} |f(x)|^2 dx\right)\left(\int_{-\infty}^{\infty} |g(x)|^2 dx\right)$$

Standard Formulation

Cauchy-Schwartz Inequality

$$\left(\int_{-\infty}^{\infty} |f(x)g(x)|dx\right)^2 \leq \left(\int_{-\infty}^{\infty} |f(x)|^2 dx\right)\left(\int_{-\infty}^{\infty} |g(x)|^2 dx\right)$$

Proof:

- Use most important inequality in math: $|z|^2 \geq 0$.
- Mechanics:

$$0 \leq \int_{-\infty}^{\infty} (af(x) + bg(x))^2 \, dx.$$

Application: $\mathbb{E}[XY]^2 \leq \mathbb{E}[X^2]\,\mathbb{E}[Y^2]$.

Application to Covariance and Correlations I

Covariance

The covariance of X and Y is defined to be

$$\mathrm{CoVar}(X, Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)].$$

Algebra yields

$$\mathrm{CoVar}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

Note that

$$\mathrm{CoVar}(X, X) = \mathrm{Var}(X).$$

Application to Covariance and Correlations II

Correlation Coefficient

Define the correlation coefficient by

$$\rho(X, Y) \ = \ \frac{\mathrm{CoVar}(X, Y)}{\sqrt{\mathrm{Var}(X) \cdot \mathrm{Var}(Y)}}.$$

Key input: $f_{X,Y}(x, y) = \sqrt{f_{X,Y}(x, y)} \cdot \sqrt{f_{X,Y}(x, y)}$.

Allows us to write $\mathbb{E}[(X - \mu_X)(Y - \mu_Y)]$ as

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)\sqrt{f_{X,Y}(x, y)} \cdot (y - \mu_Y)\sqrt{f_{X,Y}(x, y)} dxdy.$$

This is a very powerful, common technique (see also the Cramer-Rao inequality).

Clicker Questions

Generalizing collecting toys

Collecting toys

Let $H_c = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{c} \approx \log c$ be the c^{th} harmonic number. We know that if each of c prizes is equally likely each time, on average we need cH_c boxes to have each prize at least once. How many boxes are needed on average to have each prize at least twice?

Generalizing collecting toys

Collecting toys

Let $H_c = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{c} \approx \log c$ be the c^{th} harmonic number. We know that if each of c prizes is equally likely each time, on average we need cH_c boxes to have each prize at least once. How many boxes are needed on average to have each prize at least twice?

- (a) About cH_c boxes.
- (b) About $cH_{2c} \approx c\log(2c)$ boxes.
- (c) About $1.3cH_c$ boxes.
- (d) About $1.7cH_c$ boxes.
- (e) About $2cH_c$ boxes.
- (f) About $cH_c^2 \approx c\log^2 c$ boxes.x

Three Hat Problem

Problem Statement

3 mathematicians equally likely to have white or black hat. Each sees color of other hats, but not own. On three, each says ‘white’, ‘black’, or stays silent. If all speaker correct win $1 million; if even one is wrong lose $1 million. What is their winning percentage if play optimally?

Three Hat Problem

Problem Statement

3 mathematicians equally likely to have white or black hat. Each sees color of other hats, but not own. On three, each says 'white', 'black', or stays silent. If all speaker correct win $1 million; if even one is wrong lose $1 million. What is their winning percentage if play optimally?

- (a) About 12.5%
- (b) About 25%
- (c) About 37.5%
- (d) About 50%
- (e) About 62.5%
- (f) About 75%
- (g) About 87.5%
- (h) About 100%