# Math 104: Probability

Ryan Blair

University of Pennsylvania

Thursday March 14, 2013

# Outline

# Probability Density Functions

## Definition

A **probability density function** for a continuous random variable $X$ is a function $f$ defined on $(-\infty, \infty)$ such that:

1. $f$ is continuous, except possibly at finitely many points.
2. $f(x) \geq 0$ for every $x$;
3. $\int_{-\infty}^{\infty} f(x)\, dx = 1$.

The relationship of $f$ to $X$ is as follows. The probability that $X$ takes a value in the interval $[c, d]$ is given by

$$P(c \leq X \leq d) = \int_c^d f(X)\, dX.$$

# Mean, Median, Standard deviation

1. The **expected value** or **mean** of a continuous random variable $X$ with prob. density fcn. $f$ is $\mu = E(X) = \int_{-\infty}^{\infty} Xf(X)\ dX$.

# Mean, Median, Standard deviation

1. The **expected value** or **mean** of a continuous random variable $X$ with prob. density fcn. $f$ is $\mu = E(X) = \int_{-\infty}^{\infty} Xf(X)\, dX$.
2. The median of a continuous random variable $X$ with probability density function $f$ is the number $m$ for which

$$\int_{-\infty}^{m} f(X)\, dX = \frac{1}{2} \quad \text{and} \quad \int_{m}^{\infty} f(X)\, dX = \frac{1}{2}.$$

# Mean, Median, Standard deviation

1. The **expected value** or **mean** of a continuous random variable $X$ with prob. density fcn. $f$ is $\mu = E(X) = \int_{-\infty}^{\infty} Xf(X)\,dX$.
2. The median of a continuous random variable $X$ with probability density function $f$ is the number $m$ for which

$$\int_{-\infty}^{m} f(X)\,dX = \frac{1}{2} \quad \text{and} \quad \int_{m}^{\infty} f(X)\,dX = \frac{1}{2}.$$

3. The **variance** of a random variable $X$ with probability density function $f$ is the expected value of $(X-\mu)^2$ (where $\mu$ is the exp. value of $X$):

$$Var(X) = \int_{-\infty}^{\infty} (X-\mu)^2 f(X)\,dX.$$

# Mean, Median, Standard deviation

1. The **expected value** or **mean** of a continuous random variable $X$ with prob. density fcn. $f$ is $\mu = E(X) = \int_{-\infty}^{\infty} Xf(X)\ dX.$
2. The median of a continuous random variable $X$ with probability density function $f$ is the number $m$ for which

$$\int_{-\infty}^{m} f(X)\ dX = \frac{1}{2} \qquad \text{and} \qquad \int_{m}^{\infty} f(X)\ dX = \frac{1}{2}.$$

3. The **variance** of a random variable $X$ with probability density function $f$ is the expected value of $(X-\mu)^2$ (where $\mu$ is the exp. value of $X$):

$$Var(X) = \int_{-\infty}^{\infty} (X-\mu)^2 f(X)\ dX.$$

4. The **standard deviation** of $X$ is $\sigma_X = \sqrt{Var(X)}$.

# Frequently-used Distributions

**Uniform distribution**:

$$f(X) = \frac{1}{b-a} \quad \text{if} \;\; a \leq X \leq b, \qquad f(X) = 0 \quad \text{otherwise}.$$

Example: We spin an anchored arrow and let $X$ be angle (in radians) where it stops. The distribution function for $X$ is

$$f(X) = \frac{1}{2\pi} \quad \text{if} \;\; 0 \leq X < 2\pi, \qquad f(X) = 0 \quad \text{otherwise}.$$

# Normal distribution

**Normal distribution** (the bell curve):

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

The mean is $\mu$ and the standard deviation is $\sigma$.

The normal distribution is used constantly, because it reasonably fits tons of things (height, exam scores, blood dissipation, etc.).

*The mean and median of a normal distribution are equal.*

# Normal distribution, continued

Given any normal distribution

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

we have

$$P(\mu - \sigma \leq X \leq \mu + \sigma) \approx 0.68269,$$
$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx 0.95450,$$
$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) \approx 0.99730,$$
$$P(\mu - 4\sigma \leq X \leq \mu + 4\sigma) \approx 0.99994.$$

In other words, about 68% of the time the random variable $X$ lies within 1 std. deviation from the mean, about 95.5% of the time it lies within 2 std. deviations from the mean, etc.

# Example 4, $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$

A medical study of healthy people between 14 and 70 years of age modeled their systolic blood pressure $X$ using a normal distribution with mean $\mu = 119.7$ mm Hg and standard deviation $\sigma = 10.9$ mm Hg.

(a) Stage 1 high blood pressure is defined as between 140 Hg and 160 Hg. What percentage of these people has it?
(b) What percentage has blood pressure between 160 and 180 mm Hg (stage 2 high blood pressure)?
(c) What percentage has blood pressure between 90 and 120 (normal blood pressure)?

# Example 4, continued

$$(a)P(140 \le X \le 160) = \int_{140}^{160} f(X)\ dX = \int_{140}^{160} \frac{1}{\sigma\sqrt{2\pi}} e^{-(X-\mu)^2/2\sigma^2}\ dX$$
$$= \int_{140}^{160} \frac{1}{10.9 * \sqrt{2\pi}} e^{-(X-119.7)^2/2*10.9^2}\ dX$$
$$\approx 0.031165.$$

Thus, about 3.12% of the population has stage 1 high blood pressure.

$$(b)P(160 \le X \le 180) = \int_{160}^{180} f(X)\ dX \approx 0.0001090.$$

About 0.0109 percent of the population has stage 2 high blood pressure.

$$(c)P(90 \le X \le 120) = \int^{120} f(X)\ dX \approx 0.50776.$$