Continuous Random Variables

Def: A r.v. X is *continuous* if there is a non-negative function f on $\mathbb{R}^1$ having the property that for any nice/measurable set B of real numbers

$$\mathbb{P}(X \in A) = \int_A f(x)dx.$$

Taking $A = (-\infty, \infty)$, we see that $\int_{-\infty}^{\infty} f(x)dx = 1$. The function f or $f_X(x)$ is called the *probability density function* (p.d.f.) of X.
The (cumulative) distribution function $F(x)$ or $F_X(x)$ for continuous r.v. X is

$$(cdf) \qquad F(x) = \mathbb{P}(X \leq x) \quad -\infty < x < \infty$$

and $f(x) = F'(x)$.
Basic facts for continuous r.v with density $f(x)$:
- $\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx$.
- $\mathbb{P}(X = a) = \int_a^a f(x)dx = 0$.
- $\mathbb{P}(X < a) = \mathbb{P}(X \leq a)$.
- It is useful to think $f(x)dx \cong \mathbb{P}(x \leq X \leq x + dx) \cong \mathbb{P}(X \approx x)$.

EX: 6.1; 6.2.
HW: S6.1: 1, **2**, 3, 5, **8**; S6.2: 1, **2**, 5, **8**; S6.3: 1, **4**, 5, **6**; R6: 1, 3, **4**, **6**.

Density Function of a Function of a R.V.

If the density function of a random variable X is known as $f_X(x)$, how should one obtain the density function of the random variable $Y = h(X)$?

Ex 6.3: Let X be a continuous random variable with the probability density function

$$f(x) = \begin{cases} 2/x^2 & \text{if } 1 < x < 2 \\ 0 & \text{elsewhere} \end{cases}$$

Find the distribution and the density functions of $Y = X^2$.

Ex 6.4: Let X be a continuous random variable with distribution function F and probability density function f. In terms of f, find the distribution and the density functions of $Y = X^3$.

General Method for Finding Density Function of $h(X)$

Let X be a continuous r.v. having p.d.f. $f_X(x)$. For any nice/measurable function h, the r.v. $Y = h(X)$ has p.d.f

$$
\begin{aligned}
f_Y(t) &= F'_Y(t), \\
F_Y(t) &= \mathbb{P}(Y \le t) \\
&= \mathbb{P}(h(X) \le t) \\
&= \mathbb{P}(X \in B(h,t)) \\
&= \int_{B(h,t)} f_X(x)dx \\
&= \dots
\end{aligned}
$$

where $B(h,t)$ is a set determined by h and t.

• The Theorem 6.1 in the textbook is not just bad, but ugly. It is the type of mathematics that we all should avoid! Note that the theorem does not work for the simple example below.

Ex: Let X be uniform on $(-1,1)$ and $Y = X^2$. Find $f_Y(x)$.

The Inverse Transformation Method in Simulation

Theorem 13.2: Let U be a uniform $(0,1)$ random variable. For any continuous distribution function F, if we define the random variable Y by

$$Y = F^{-1}(U)$$

then the r.v. Y has distribution function F, where $F^{-1}(x)$ is defined to equal that value y for which $F(y) = x$.

Pf: Details in class.

Ex: Exponential r.v. with parameter $\lambda > 0$, i.e. $f(x) = \lambda e^{-\lambda x}$ for $x > 0$.

Ex: The Cauchy r.v. with parameter $\theta \in \mathbb{R}$, i.e. the p.d.f. is

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{1 + (x - \theta)^2}, \qquad -\infty < x < \infty.$$

Expectation for continuous r.v.

Def: $\mathbb{E}\,X = \int_{-\infty}^{\infty} x f(x) dx.$

Theorem 6.3: If X is a continuous r.v. with p.d.f. $f(x)$, then for any function $h(x)$,

$$\begin{aligned} \mathbb{E}\,h(X) &= \int_{-\infty}^{\infty} h(x) f(x) dx \\ &= \text{integral of value} \times \text{probability (density)} \end{aligned}$$

Ex: If the p.d.f of r.v. X is given by

$$f(x) = \begin{cases} C(4x - 2x^2) & 0 < x < 2, \\ 0 & \text{otherwise} \end{cases}$$

Find

1. The value of C,
2. $\mathbb{P}(X > 1)$
3. $\mathbb{E}\,X$
4. $\text{Var}(X)$

Ex: 6.8; 6.9; 6.10; 6.11.

The Uniform Random Variable

A r.v X is *uniform* on the interval (a, b) if its p.d.f is given by

$$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b, \\ 0 & \text{otherwise} \end{cases}$$

- The expectation $\mathbb{E}\, X = \int_a^b x \cdot \frac{1}{b-a} dx = (a+b)/2$.
- The variance $\mathrm{Var}(X) = (b-a)^2/12$.

Ex: Buses arrive at a specific stop at 15-minute intervals starting at 7 A.M. If a passenger arrives at the stop at a time that is uniformly distributed between 7 and $7:30$, find the probability that he waits
(a). less than 5 minutes for a bus;
(b). more than 10 minutes for a bus.

Ex: A stick of length 1 is split at a point U that is uniformly distributed over $(0, 1)$. Determine the expected length of the piece that contains the point p, $0 \le p \le 1$.
EX: 7.1; 7.2; 7.3.
HW: S7.1: 1, 3, **4**, **6**; S7.2: 1, **6**, 7, 9, **10**; S7.3: **2**, 3, **6**, 7; R7: 1, **2**, 3, 9, **14**.

Normal Random Variables

A r.v X is *normal* with parameter μ and σ^2, $\sigma > 0$, if its p.d.f is given by

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

• We will show in class that $\int_{-\infty}^{\infty} f(x)dx = 1$.
• The expectation $\mathbb{E}\,X = \mu$.
• The variance Var$(X) = \sigma^2$ and SD$(X) = \sigma$.
• We use $X \sim N(\mu, \sigma^2)$ to denote normal r.v. with mean μ and variance σ^2.
• When $\mu = 0$ and $\sigma = 1$, $N(0,1)$ is called standard normal and denoted by Z with c.d.f

$$\Phi(x) = \mathbb{P}(Z \le x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$$

and $\Phi(-x) = 1 - \Phi(x)$.
• Standardization: If $X \sim N(\mu, \sigma^2)$, then

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1).$$

The Normal Approximation to the Binomial Distribution

The DeMoivre-Laplace limit theorem: Let S_n be the number of successes in n independent Bernoulli trials with $\mathbb{P}(\text{Success}) = p$. Then for any $a < b$,

$$\begin{aligned} \mathbb{P}\left(a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b \right) &\rightarrow \mathbb{P}(a \leq Z \leq b) \\ &= \Phi(b) - \Phi(a) \end{aligned}$$

as $n \rightarrow \infty$.

Correction for Continuity:

$$\begin{aligned} \mathbb{P}(i \leq S_n \leq j) &= \mathbb{P}(i - 1/2 \leq S_n \leq j + 1/2) \\ \mathbb{P}(S_n = k) &= \mathbb{P}(k - 1/2 \leq S_n \leq k + 1/2). \end{aligned}$$

Ex: Let X be the number of times that a fair coin, flipped 100 times, lands heads. Find an approximation of the probability that $50 < X \leq 55$.

EX: 7.4; 7.5; 7.6; 7.7.

Exponential Random Variable

A continuous r.v X is *exponential* with parameter $\lambda > 0$ if its p.d.f is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & \text{otherwise} \end{cases}$$

- The expectation $\mathbb{E}\, X = \int_0^\infty x\lambda e^{-\lambda x} dx = 1/\lambda$.
- The variance $\mathrm{Var}(X) = 1/\lambda^2$.
- The *memoryless* property: For all $s, t \geq 0$,

$$\mathbb{P}(X > s + t \,|\, X > t) = \mathbb{P}(X > s).$$

Hazard Rate Functions: Consider a positive continuous r.v. X what we interpreted as being the lifetime of some item, having c.d.f. $F(x)$ and p.d.f. $f(x)$. The *hazard rate* or *failure rate* function $\lambda(t)$ of F is $\lambda(t) = \frac{f(t)}{1-F(t)}$ which represent the conditional probability intensity that a t-unit-old item will fail, i.e.

$$\mathbb{P}(X \in (t, t + dt)|X > t) \approx \lambda(t)dt.$$

For exponential r.v's, the hazard rate $\lambda(t) = \lambda$.
EX: 7.10; 7.11; 7.12.

Joint Distributions of Two R.V's

Def: For any two r.v.'s X and Y, the *joint cumulative probability distribution function* of X and Y is defined by

$$F_{X,Y}(a,b) = \mathbb{P}(X \le a,\, Y \le b), \quad -\infty < a, b < \infty$$

The distribution of X, also called marginal, can be obtained from the joint one,

$$F_X(a) = \mathbb{P}(X \le a) = F(a, \infty)$$

If both X and Y are discrete, then *joint probability mass function* is defined by

$$p_{X,Y}(x,y) = \mathbb{P}(X = x,\, Y = y).$$

Thus

$$p_X(x) = \mathbb{P}(X = x) = \sum_y p_{X,Y}(x,y),$$
$$p_Y(y) = \mathbb{P}(Y = y) = \sum_x p_{X,Y}(x,y)$$

and they are called, respectively, the *marginal probability mass functions* of X and Y.

HW: S8.1: 1, **2**, 5, 9, **10**; S8.2: 1, 3, **12**, 15, **20**; S8.3: 1, **2**, 7, **8**: S8.4: 1, **4**, 5; R8: 1, **2**, **6**, 7, 15.

Ex: Toss a fair coin three times.

$$\begin{aligned} X &= \text{\# of heads on the first toss} \\ Y &= \text{total \# of heads} \end{aligned}$$

EX: 8.1; 8.2; 8.3.

We have for expectations,

$$\mathbb{E}\,X = \sum_x x p_X(x), \quad \mathbb{E}\,Y = \sum_y y p_Y(y),$$
$$\mathbb{E}\,h(X,Y) = \sum_x \sum_y h(x,y) p_{X,Y}(x,y).$$

In particular, for any random variables X and Y,

$$\mathbb{E}\,(X+Y) = \mathbb{E}\,X + \mathbb{E}\,Y.$$

Jointly Continuous Random Variables

The non-negative function $f(x,y)$ is called *joint probability density function* (j.p.d.f) of X and Y if for any nice/measurable region $R \subset \mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$

$$\mathbb{P}((X,Y) \in R) = \int \int_R f(x,y) dx dy.$$

- Taking $R = \mathbb{R}^2$, then $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) dx dy = 1$.
- Taking $R = \{(x,y) : x \in A, y \in B\}$, then

$$\mathbb{P}(X \in A, Y \in B) = \int_A \int_B f(x,y) dy dx.$$

- Connection between j.c.d.f. and j.p.d.f.

$$F_{X,Y}(a,b) = \mathbb{P}(X \leq a, Y \leq b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f(x,y) dy dx,$$

and $f(x,y) = \dfrac{\partial^2}{\partial a \partial b} F(a,b)$.

- Marginal p.d.f of X:

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y) dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f(x,y) dx$$

- It is useful to think

$$f(x,y) dx dy \cong \mathbb{P}(x \leq X \leq x + dx, y \leq Y \leq y + dy) \cong \mathbb{P}(X \approx x, Y \approx y).$$

Ex: Given

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & 0 < x < \infty, 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Compute
1. $\mathbb{P}(X > 1,\ Y < 1)$,
2. $\mathbb{P}(Y < X)$
3. $\mathbb{P}(Y < a)$

Ex: For uniform distribution on a circle of radius R, the j.p.d.f is

$$f(x,y) = \begin{cases} c & x^2 + y^2 \leq R^2 \\ 0 & \text{otherwise} \end{cases}$$

Find
1. the value for c,
2. $F_X(x)$ and $f_Y(y)$,
3. $\mathbb{P}(D < a)$ where $D = \sqrt{X^2 + Y^2}$,
4. $\mathbb{E}\,D$.
5. Intuitive explanation for the formula of $\mathbb{E}\,D$.

Ex: 8.4; 8.6; 8.7.

We have for expectations,

$$\mathbb{E}\,X = \int_{-\infty}^{\infty} x f_X(x)dx, \quad \mathbb{E}\,Y = \int_{-\infty}^{\infty} y f_Y(y)dy,$$

$$\mathbb{E}\,h(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y) f_{X,Y}(x,y)dxdy.$$

Independent Random Variables

The r.v's X and Y are *independent* if for any two nice/measureable sets of real numbers A and B,

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B).$$

i.e., X and Y are independent if for all nice/measureable A and B, the events $E_A = \{X \in A\}$ and $F_B = \{Y \in B\}$ are independent.

- X and Y are ind. iff for all $a, b \in \mathbb{R}$,

$$\mathbb{P}(X \leq a, Y \leq b) = \mathbb{P}(X \leq a)\mathbb{P}(Y \leq b)$$

i.e. $F_{X,Y}(a,b) = F_X(a)F_Y(b)$ for all $a, b \in \mathbb{R}$.

- X and Y are ind. iff for all $x, y \in \mathbb{R}$,

$$f_{X,Y}(x,y) = f_X(x)f_Y(y) \text{ or } p(x,y) = p_X(x)p_Y(y).$$

- X and Y are ind. iff their joint p.d.f can be expressed as

$$f_{X,Y}(x,y) = h(x)g(y), \quad -\infty < x, y < \infty.$$

Ex: Let X and Y be ind. and uniform on $(0, 60)$. Find $\mathbb{P}(|X - Y| > 10)$.

Theorem: Let X and Y be independent random variables and $g : \mathbb{R} \to \mathbb{R}$ and $h : \mathbb{R} \to \mathbb{R}$ be real-valued functions. Then $g(X)$ and $h(Y)$ are independent random variables and

$$\mathbb{E}\,(g(X)h(Y)) = \mathbb{E}\,g(X) \cdot \mathbb{E}\,h(Y).$$

Pf: Details in class.

Ex: If X and Y are independent, then

$$\begin{aligned} \mathbb{E}\,(XY) &= \mathbb{E}\,X \cdot \mathbb{E}\,Y \\ \mathbb{E}\,(XY^2) &= \mathbb{E}\,X \cdot \mathbb{E}\,Y^2 \\ \mathbb{E}\,\Big((X-1)(Y^2+1)\Big) &= \mathbb{E}\,(X-1) \cdot \mathbb{E}\,(Y^2+1) \end{aligned}$$

Ex: 8.12; 8.14; 8.15.

Buffon's Needle Problem

A table is rued with equidistant parallel lines a distance D apart. A needle of length L, where $L \leq D$, is randomly thrown on the table. What is the probability that the needle will intersect one the lines (the other possibility being that the needle will be completely contained in the strip between two lines)?

Sol. 1: The probability is $2L/(D\pi)$ via calculus as given in the book.

Sol. 2: Find the expected number of crossings for all L and D via linearity of expectation, without any calculation!

- There are some simulations on the Web.
- It's possible to get excellent results from the needle-tossing technique, if you don't mind cheating a bit as you do it. See, for example, in history/HistTopics/ Pi_through_the_ages.html, which includes this:

“Various people have tried to calculate π by throwing needles. The most remarkable result was that of Lazzerini (1901), who made 34080 tosses and got $\pi = 355/113 = 3.1415929$ which, incidentally, is the value found by Tsu Ch’ung Chi. This outcome is suspiciously good, and the game is given away by the strange number 34080 of tosses. Kendall and Moran comment that a good value can be obtained by stopping the experiment at an optimal moment. If you set in advance how many throws there are to be then this is a very inaccurate way of computing π. Kendall and Moran comment that you would do better to cut out a large circle of wood and use a tape measure to find its circumference and diameter”.

“Still on the theme of phoney experiments, Gridgeman, in a paper which pours scorn on Lazzerini and others, created some amusement by using a needle of carefully chosen length $L = 0.7857$, throwing it twice, and hitting a line once. His estimate for π was thus given by $2 \times 0.7857/\pi = 1/2$ from which he got the highly creditable value of $\pi \approx 3.1428$. He was not being serious!”

Revisit Buffon Needle Problem

Here is the most beautiful proof of the problem by extending it! Instead of calculating the probability, calculate the expected number of crossings. With $L < D$, that's the same thing, but we can get the expected number without that restriction. We use the fact that the expected value of a sum is the sum of the expected values. This holds even without independence. Imagine the needle divided into pieces. The expected number of crossings of the whole needle is the sum of the expected numbers for the parts. If the parts are identical, they have the same expected number, so the expected number is proportional to the length of the needle. We want to know the constant of proportionality. The argument even applies to a curved needle (or even a bootlace, or a handful of needles), so imagine a needle bent into a circle of length $L = \pi D$. (This is greater than D, but that doesn't matter as stated above). Such a needle always has exactly two crossings apart from when it is tangential to two lines, an event that has zero probability. The expected number is therefore 2 giving a constant of proportionality of $2/(\pi D)$. So for a needle of length L we get $2L/(\pi D)$.

Conditional Distributions

The *conditional* probability mass function of X given $Y = y$ is

$$\mathbb{P}_{X|Y}(x|y) = \mathbb{P}(X = x \,|\, Y = y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

The *conditional* probability density function of X given $Y = y$ is

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$$

and the conditional c.d.f is

$$F_{X|Y}(x|y) = \mathbb{P}(X \le x \,|\, Y = y) = \int_{-\infty}^{x} f_{X|Y}(t|y)dt.$$

- If X and Y are ind., then $f_{X|Y}(x|y) = f_X(x)$.

- Geometric meaning.

- $f_{X|Y}(x|y)dx \approx \mathbb{P}(x \le X \le x + dx | y \le Y \le y + dy)$.

EX 8.22: $Y \sim$ uniform$(0, 1)$ and $X \sim$ uniform$(0, Y)$ Find $f_X(x)$ and $\mathbb{E}\, X$.

Conditional Expectation

The *conditional expectation* of X given $Y = y$ is

$$\mathbb{E}(X|Y=y) \;\; = \sum_x x \cdot \mathbb{P}(X = x|Y = y)$$
$$\left(= \int_{-\infty}^{\infty} x \cdot f_{X|Y}(x|y)dx\right)$$

- $\mathbb{E}(X|Y = y)$ is a function of y.
- The conditional expectations are simply ordinary expectations computed relative to conditional distributions.
- $\mathbb{E}(h(X)|Y = y) = \sum_x h(x)p_{X|Y}(x|y)$ in discrete case, and

$\mathbb{E}(h(X)|Y = y) = \int_{-\infty}^{\infty} h(x)f_{X|Y}(x|y)dx$ in continuous case.

EX 8.24: Given

$$f(x,y) = \begin{cases} e^{-y} & \text{if } y > 0,\ 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $\mathbb{E}(X|Y = 2)$.

Joint Prob. Distribution of Functions of R.V's

Let X and Y be r.v's with j.p.d.f $f_{X,Y}(x,y)$. Suppose that $U = h_1(X,Y)$ and $V = h_2(X,Y)$. Then we can find the j.p.d.f $g_{U,V}(u,v)$ of (U,V) via

$$g_{U,V}(u,v) = \frac{\partial^2}{\partial u \partial v} F_{U,V}(u,v)$$

and

$$\begin{aligned} F_{U,V}(u,v) &= \mathbb{P}(U \leq u, V \leq v) \\ &= \mathbb{P}(h_1(X,Y) \leq u, h_2(X,Y) \leq v) \\ &= \int\int_{h_1(x,y)\leq u, h_2(x,y)\leq v} f_{X,Y}(x,y)dxdy \\ &= \ldots \end{aligned}$$

Ex 8.27: Let U_1, U_2 be ind. uniform r.v's on $(0,1)$. Then

$$\begin{aligned} X_1 &= \sqrt{-2\log U_1}\cos(2\pi U_2) \\ X_2 &= \sqrt{-2\log U_1}\sin(2\pi U_2) \end{aligned}$$

are ind. standard normal r.v's.

• This provides a way to generate ind. standard normal r.v's from ind. uniforms.

Sum of Independent Random Variables

If X and Y are independent, then

$$f_{X+Y}(t) = \int_{-\infty}^{\infty} f_X(t-y) f_Y(y) dy = (f_X * f_Y)(t)$$

which is the *convolution* of two density functions. Note that $(f_X * f_Y)(t) = (f_Y * f_X)(t)$.

Basic facts for sum of ind. r.v.'s X and Y:

- If $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$, then $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

- If $X \sim$ Poisson(λ_1) and $Y \sim$ Poisson(λ_2), then $X+Y \sim$ Poisson($\lambda_1 + \lambda_2$).

- If $X \sim$ binomial(n, p) and $Y \sim$ binomial(m, p), then $X + Y \sim$ binomial($n + m, p$).

Covariance

The *covariance* between X and Y, denoted by Cov(X, Y), is defined by

$$\begin{aligned} \text{Cov}(X,Y) &= \mathbb{E}\left((X - \mathbb{E}\,X)(Y - \mathbb{E}\,Y)\right) \\ &= \mathbb{E}\,(XY) - \mathbb{E}\,X \cdot \mathbb{E}\,Y. \end{aligned}$$

• Cov$(X,Y) =$ Cov(Y,X), Cov$(X,X) =$ Var(X), and for any number $a \in \mathbb{R}$, Cov$(aX,Y) = a$Cov(X,Y).
• Cov$\left(\sum_{i=1}^n X_i, \sum_{j=1}^m Y_j\right) = \sum_{i=1}^n \sum_{j=1}^m$ Cov(X_i, Y_j).
• Var$\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n$ Var$(X_i) + 2\sum_{i<j}$ Cov(X_i, X_j).
• If X_i, $1 \le i \le n$, are i.i.d, then

$$\text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i).$$

• If Cov$(X,Y) > 0$, then X and Y are positive correlated. If Cov$(X,Y) < 0$, then X and Y are negative correlated. If Cov$(X,Y) =$ 0, then X and Y are uncorrelated.
Ex: 10.12.
HW: S10.2: 3, 7, **10**, **12**; S10.3: 1, **2**, 3; R10: 7, 9, **12**.

Correlation

Let X and Y be two random variables with $0 < \sigma_X^2 = \text{Var}(X) < \infty$ and $0 < \sigma_Y^2 = \text{Var}(Y) < \infty$. The covariance between the standardized X and the standardized Y is called the *correlation coefficient* between X and Y and is denoted by

$$\begin{aligned}\rho = \rho(X,Y) &= \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y} \\ &= \frac{\mathbb{E}\,((X - \mathbb{E}\,X)(Y - \mathbb{E}\,Y))}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}}\end{aligned}$$

- $-1 \le \rho(X,Y) \le 1$.
- $\rho(X,Y) = 1$ iff $Y = aX + b$ for some constant a, b, $a > 0$.
- $\rho(X,Y) = -1$ iff $Y = aX + b$ for some constant a, b, $a < 0$.
- X and Y are linearly related iff $\rho(X,Y) = \pm 1$.

EX: 10.17; 10.18.