Stat310

Bivariate transformations

Hadley Wickham

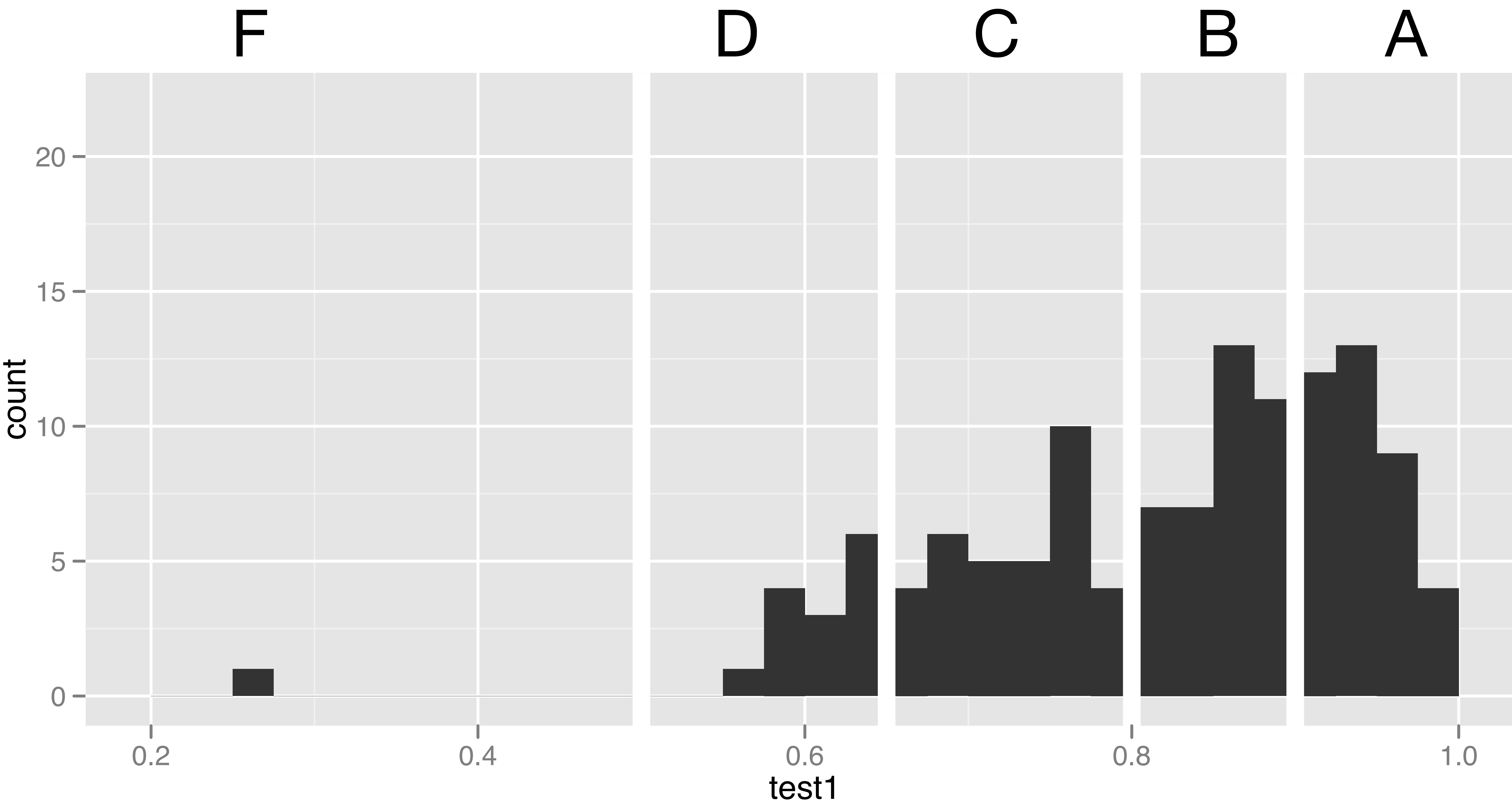
1. Test feedback
2. Correlation
3. Transformations

Test

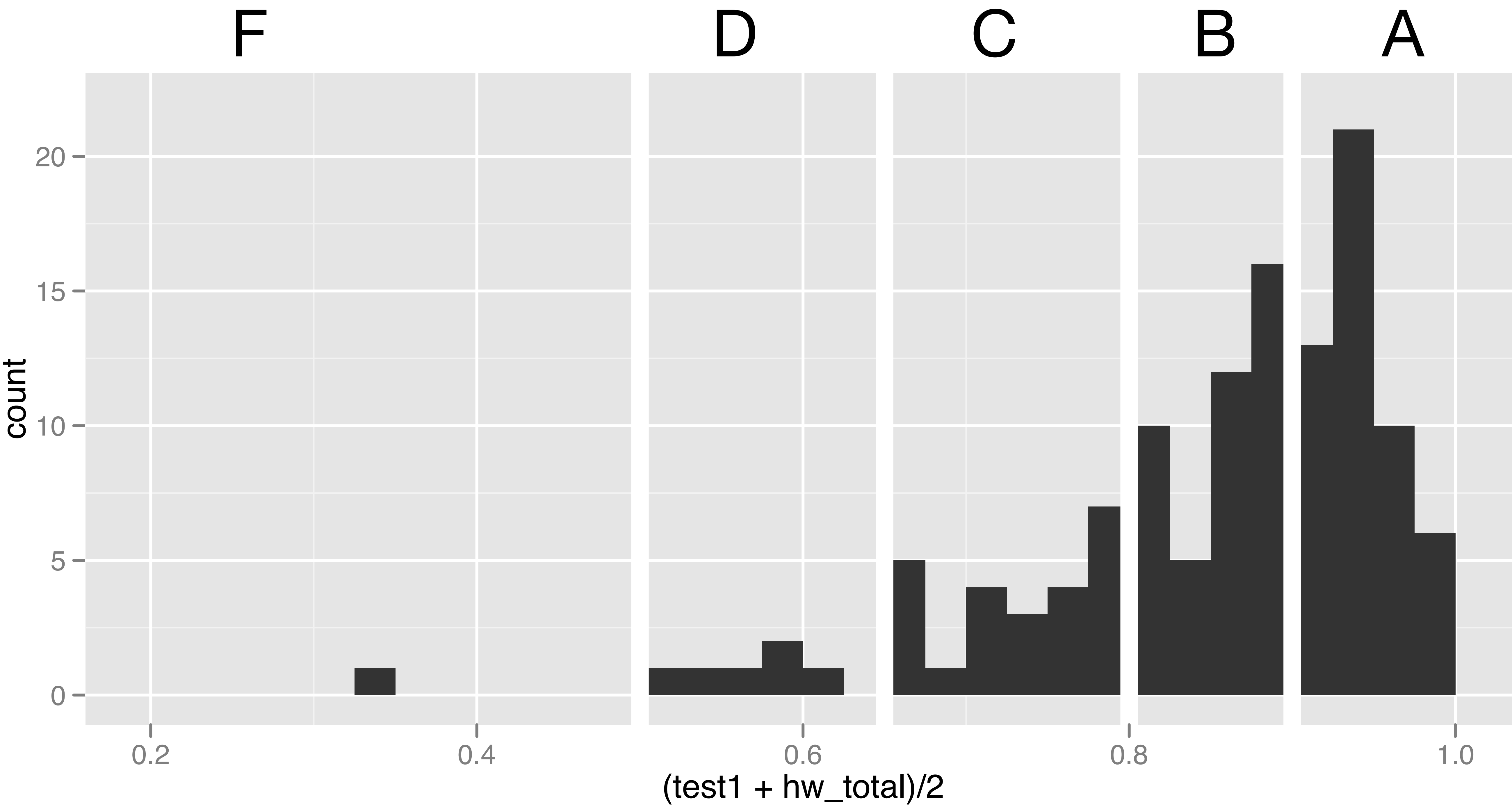
Overall

Test went went. No curving.

Model answers available online.



3 perfect scores



(Approximation - doesn’t include dropping lowest grades)

Questions

1 was a bit tricky, but if you pay attention in class, you should've been ok.

Many of you forgot the pdf conditions for 2.

A lot of you missed 3(d), despite numerous hints in class. Did well on 3(c) considering we never explicitly talked about it in class.

4 should have been easy.

Homework

Available on the website now.

Homework

Available on the website now.

But since it's late it will be bonus points. Since it's short it will count as half a homework.

Homework

Available on the website now.

But since it's late it will be bonus points. Since it's short it will count as half a homework.

(Don't forget to turn in your stats in practice essay on Thursday.)

Correlation

Independence

$P(X \mid Y) = X$

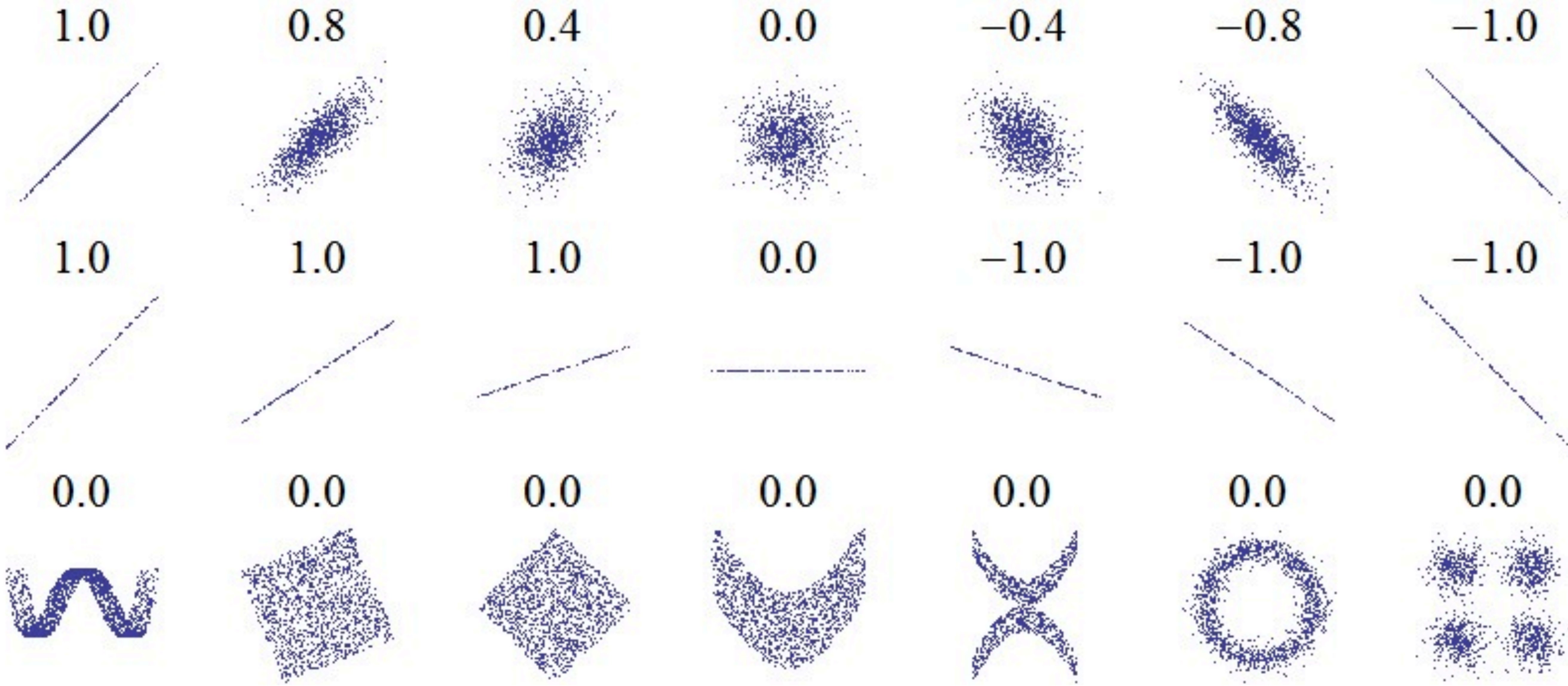
$P(Y \mid X) = Y$

$\Rightarrow f(x, y) = f(x)\, f(y),\ S = S_x \times S_y$

Dependence

Only one way for rv's to be independent.

Many ways to be dependent. Useful to have some measurements to summarise common forms of dependence: we'll use **correlation**



From: http://en.wikipedia.org/wiki/File:Correlation_examples.png

ρ

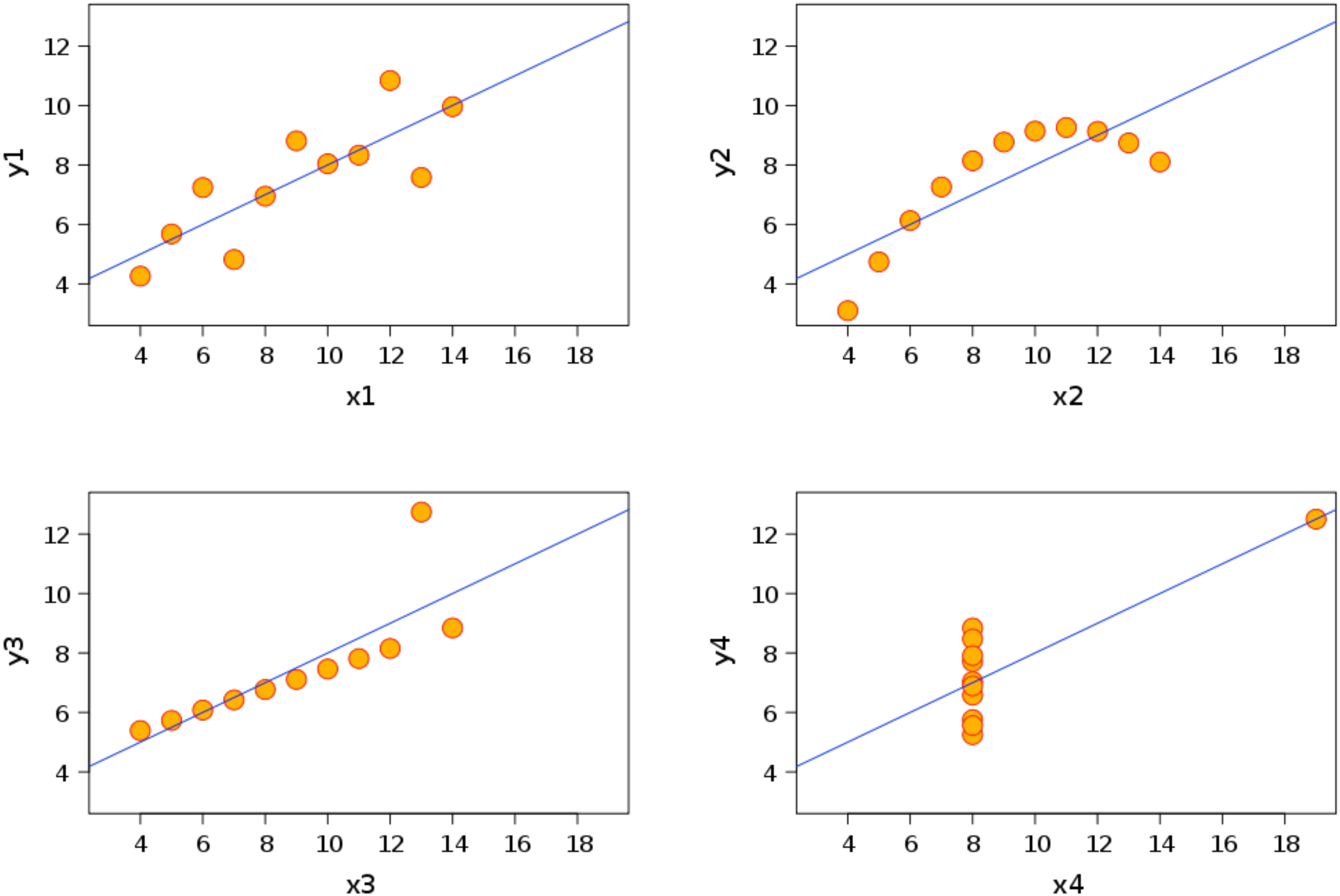
What is the range of ρ?

For what value of ρ are X and Y most strongly positively related?

For what value of ρ are X and Y most strongly negatively related?

For what value of ρ is there no (linear) relationship between X and Y?

Correlation = 0.8



Covariance

Easiest to define correlation in terms of another function: co**variance**

$Cov(X, Y) = E[\ (X - E(X))(Y - E(Y))\ ] = \sigma_{XY}$

$Cov(X, X) = Var(X) = \sigma_{XX} = \sigma_X^2$

Covariance

If $\sigma_{XY} > 0$, then X tends to increase when Y increases

If $\sigma_{XY} < 0$, then X tends to decrease when Y increases

If $\sigma_{XY} = 0$, then there is no linear relationship between X and Y
(but there may be a non-linear relationship!)

Alternative

Is there another way to compute the covariance? (Think about the two ways of computing the variance)

If X and Y are independent,
what is Cov(X, Y)?

Correlation

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

$$\rho_{XY} = \frac{\sigma_{XY}}{\sqrt{\sigma_{XX} \sigma_{YY}}}$$

Counterexample

Let X be a random variable with $E(X) = 0$, $E(X^2) = 10$, $E(X^3) = 0$, $E(X^4) = 4$. Let $Y = X^2$.

Are X and Y independent?

Are X and Y uncorrelated?

Bivariate transformation

Example

Oscar has a bad gambling problem. Every night on the way home from work he takes the X hundred dollars he earned at work that day and goes to the local casino. Oscar never wins any money but eventually stops playing to return home with Y hundred dollars.

$$f(x,y) = \frac{1}{8} \quad 0 < y < x < 4$$

$$f(x) = \frac{x}{8} \quad 0 < x < 4$$

$$f(y) = \frac{1}{8}(4 - y) \quad 0 < y < 4$$

Questions

Are X and Y independent?

Let A = (X - Y)/X and B = X. Are A and B independent? How can we tell?

Need to find pdf of A and B. Make two new random variables from two existing random variables: **transformation**

1d transformations

Distribution function technique

(always works, but hard)

Change of variable technique

(easy, but doesn't always work)

2d transformations

Distribution function technique

(hardly ever works, but easy)

Change of variable technique

(usually works, moderately hard)

Distribution function technique

CDF

As we've seen, the cdf isn't that useful for bivariate random variables. Thus, the distribution function technique isn't that useful.

However, there are a few cases where it makes the transformation very easy.

Your turn

Fill in the blanks proof to find first the pdf of max(X, Y) if X and Y are uniform, and then the pdf for any distribution of X and Y.

Change of variables technique

1d change of variables

$$\mathrm{A} = \mathrm{u}(\mathrm{X})$$

$$\mathrm{X} = \mathrm{v}(\mathrm{A})$$

$$J = \frac{d}{da}x$$

$$f_A(a) = f_X(v(a))|J|$$

2d change of variables

$$A = u_1(X, Y) \quad X = v_1(A, B)$$

$$B = u_2(X, Y) \quad Y = v_2(A, B)$$

$$f_{A,B} = f_{X,Y}(v_1(A, B), v_2(A, B))|J|$$

Be careful of two different uses of |

$$J = \left| \begin{array}{cc} \frac{\delta x}{\delta a} & \frac{\delta x}{\delta b} \\ \\ \frac{\delta y}{\delta a} & \frac{\delta y}{\delta b} \end{array} \right|$$

$$\left| \begin{array}{cc} a & b \\ c & d \end{array} \right| = ad - bc$$

Steps

Write down u_1, u_2, $f_{X,Y}$

Figure out bounds of A and B

Figure out v_1 and v_2

Compute partial derivatives

Plug into formula

Back to the problem

- Are (X - Y)/X and X independent?
- What are A and B?
- What are their ranges?
- What are u_1 and u_2?
- What are v_1 and v_2?
- What is $f_{A,B}(a, b)$?

$$f(x,y) = \frac{1}{8} \quad 0 < y < x < 4$$

$$f(x) = \frac{x}{8} \quad 0 < x < 4$$

$$f(y) = \frac{1}{8}(4 - y) \quad 0 < y < 4$$

In general

- Figuring out v_1 and v_2 and their partial derivatives is usually easy (in this class)
- Figuring out the bounds of integration is often hard! (So in a test I would give to you, or give step by step)

Next time

- More practice!