Stat310

CLTs & review

Hadley Wickham

1. MGF facts
2. Two central limit theorems
3. Review

MGF revision

mgf

The **moment generating function** (**mgf**) is $M_X(t) = E(e^{Xt})$
(Provided it is finite in a neighbourhood of 0)

Why is it called the mgf?

If $M_X(t) = M_Y(t)$ then X and Y have the same pdf/pmf.

Your turn

If X and Y are independent, what is the MGF of $Z = X + Y$? What is the MGF of $A = bX$?

Generally, for independent random variables, how do you think the mgf of the sum will be related to the individual mgfs? What about for iid random variables?

$$S_n = \sum_{i=1}^{n} X_i$$

X_i are independent

$$M_{S_n}(t) = \prod_{i=1}^{n} M_{X_i}(t)$$

X_i are iid

$$M_{S_n}(t) = (M_X(t))^n$$

CLT

Your turn

$X_1, X_2, \ldots$ are iid $N(\mu, \sigma^2)$

$$S_n = \sum_1^n X_i \qquad \bar{X}_n = \frac{S_n}{n}$$

Find their mgfs. What do you notice?

Hint: $M_X(t) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$

Your turn

Fill in the blanks to prove the central limit theorem for all random variables with a mgf.

Continuity correction

If X is discrete, we can still use the CLT. But we can make it a little better with a continuity correction.

For example, X ~ Binomial(n, p). If n is big, then we can approximate X with the normal distribution Y

Y ~~ Normal(?, ?)

$P(X \leq x) = P(X < x + 1)$, but

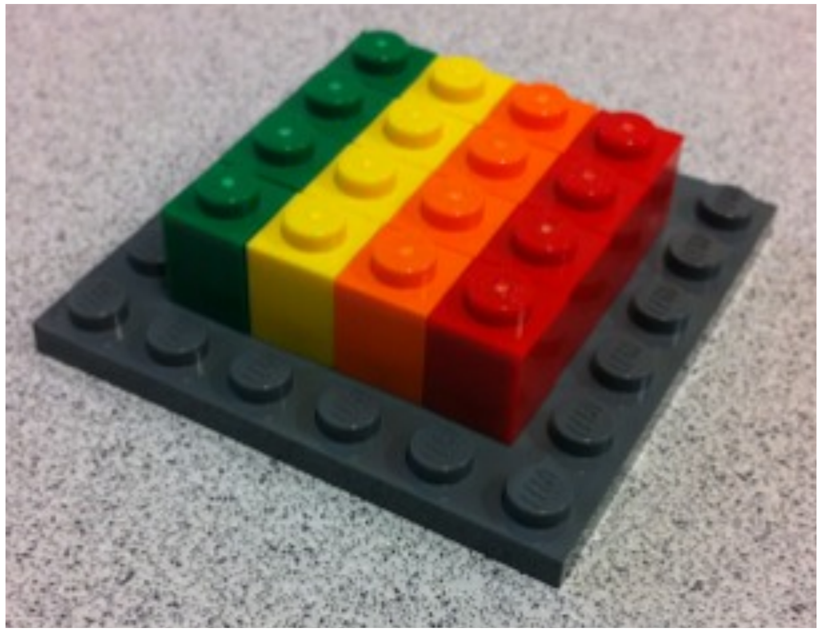
$P(Y \leq x) = P(Y \leq x + 1/2)$

Review

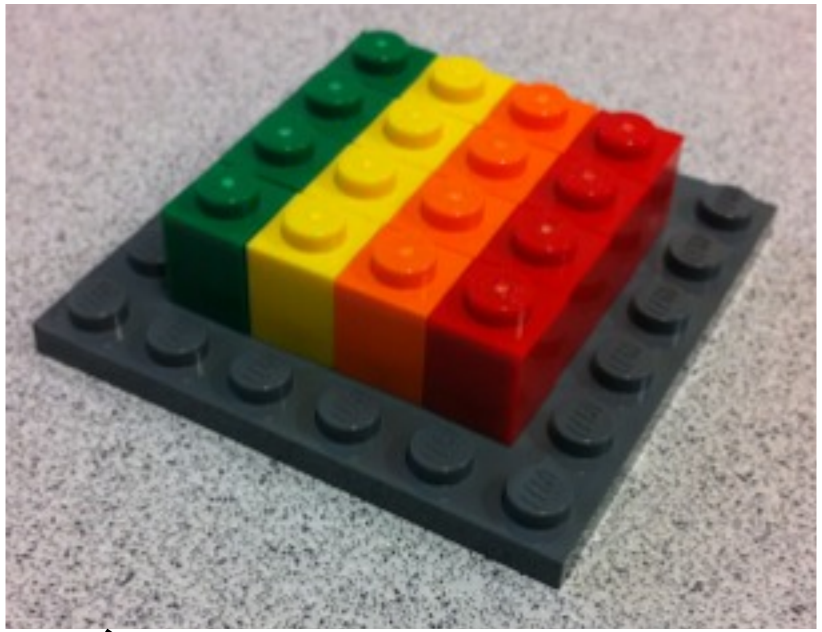
1. Bivariate distributions
2. Transformations: univariate & bivariate
3. Sequences of random variables

Bivariate distributions

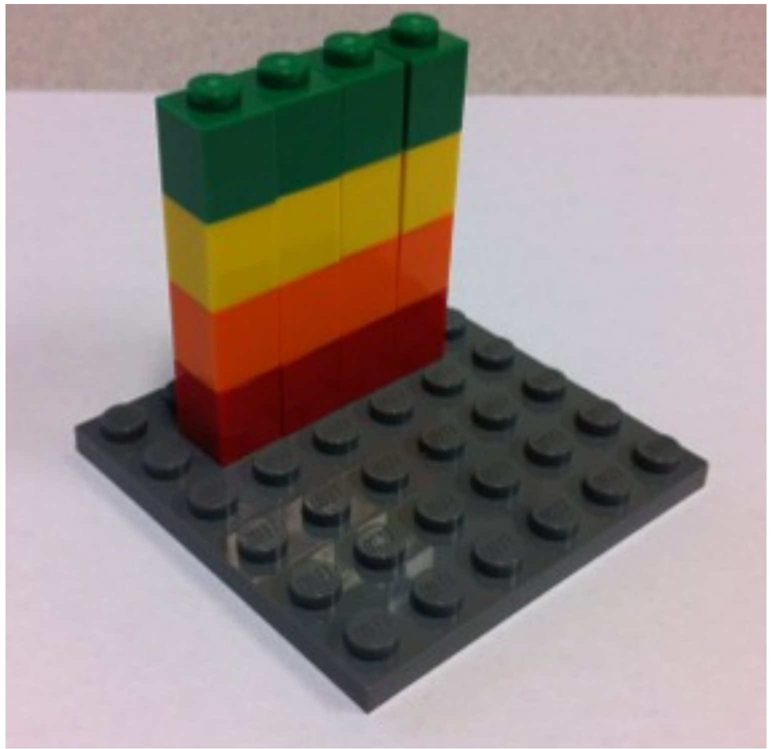
f(x, y)



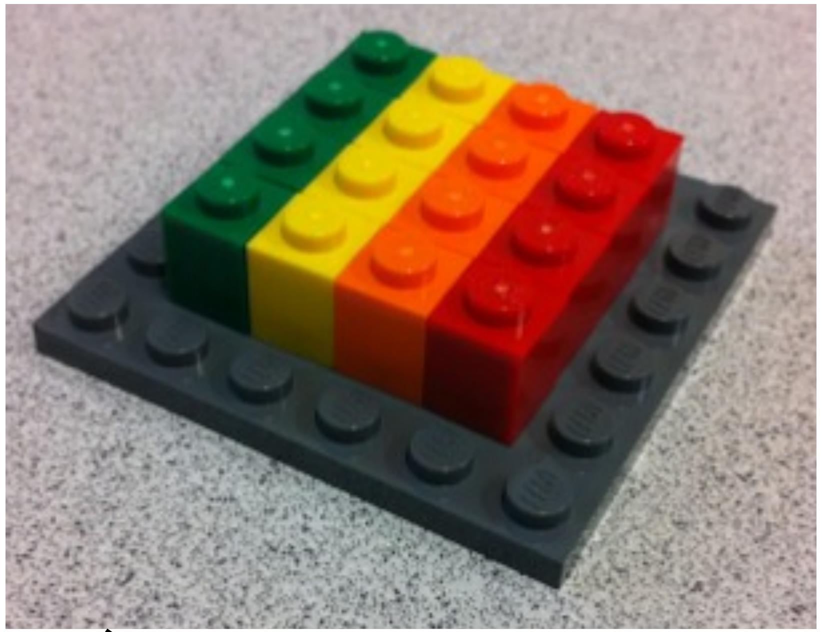
f(x, y)



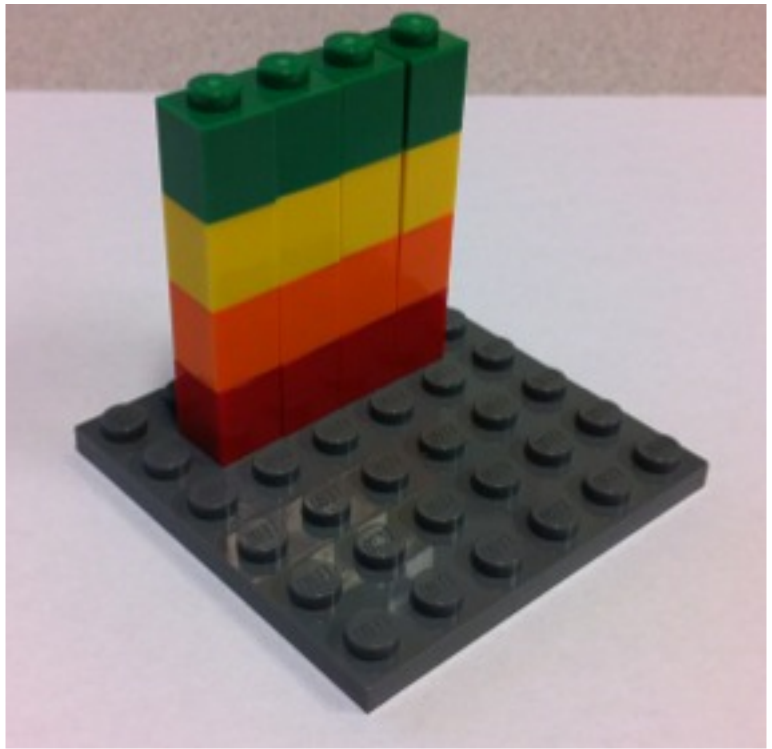
f(x)



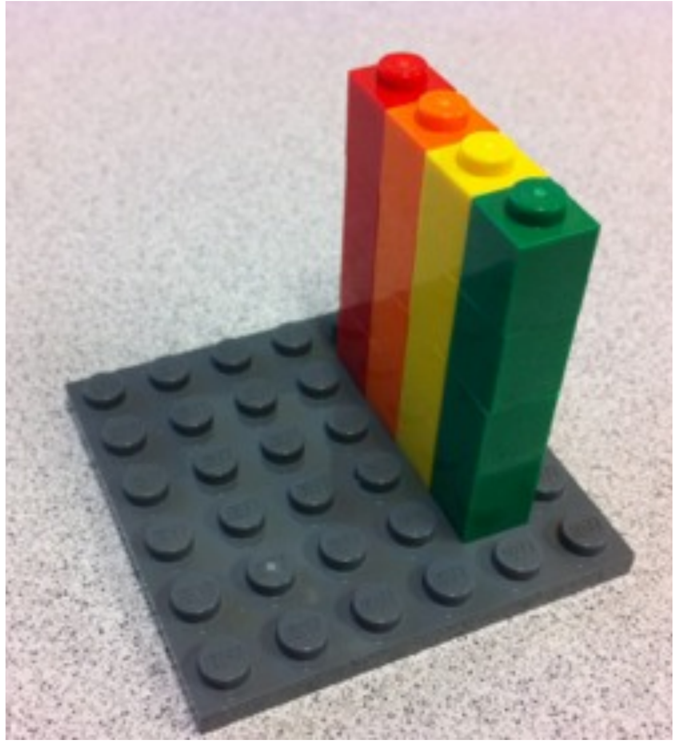
f(x, y)



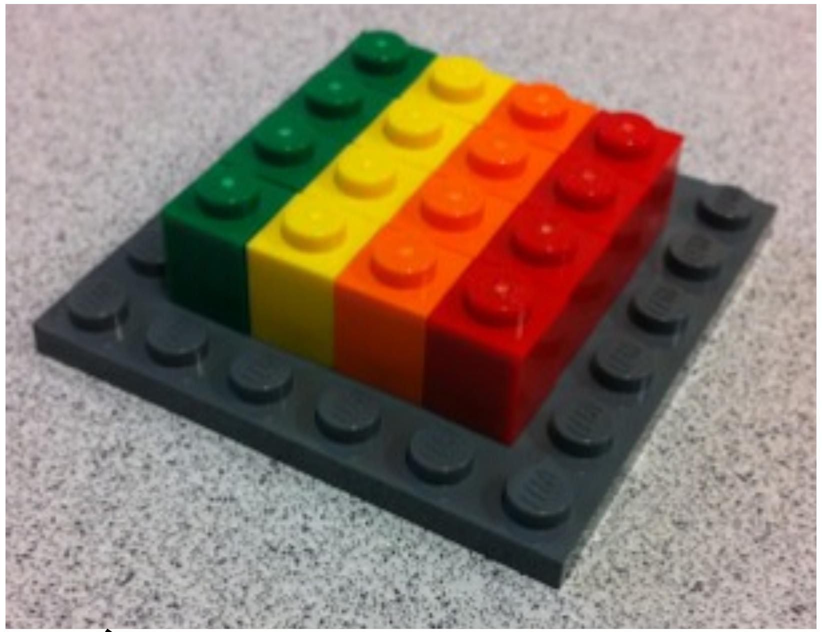
f(x)



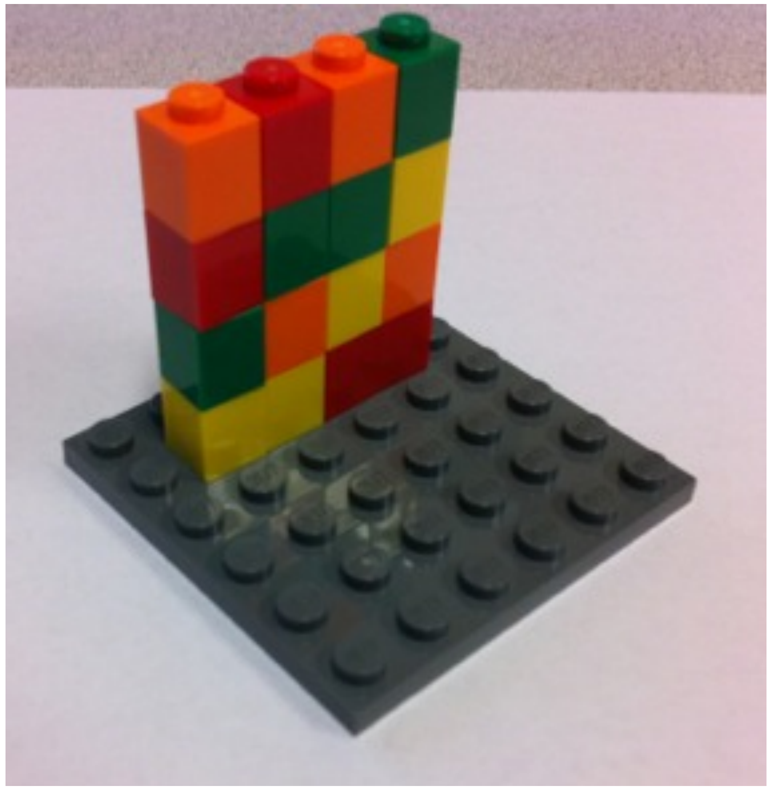
f(y)



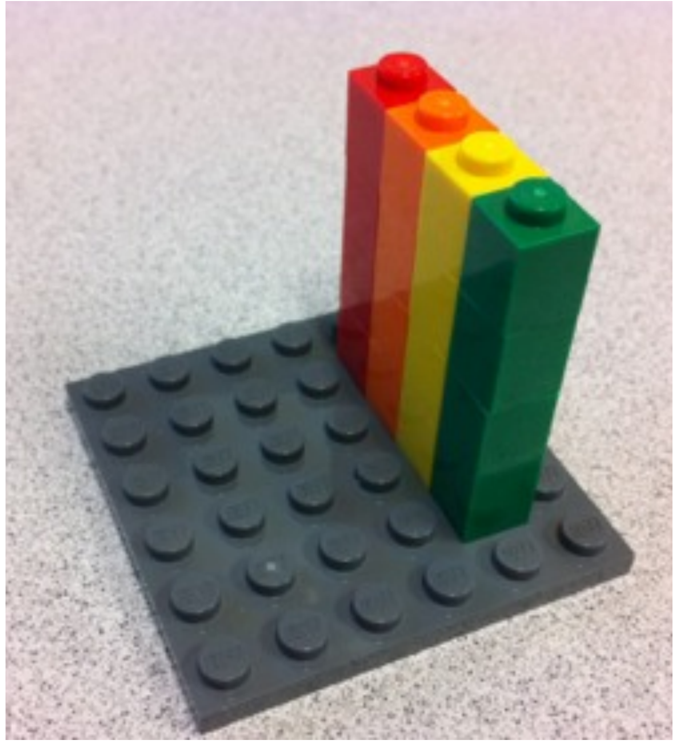
f(x, y)

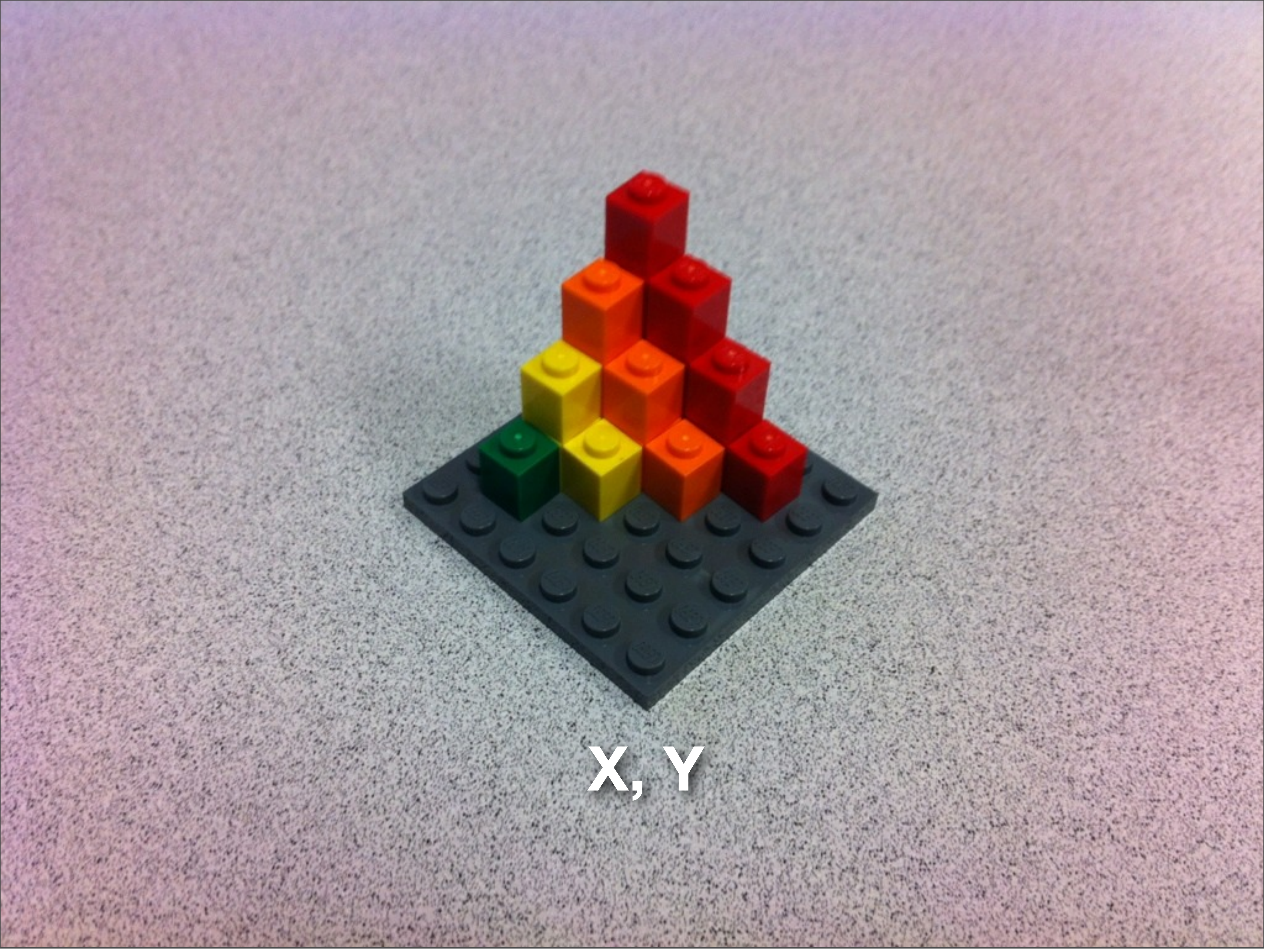


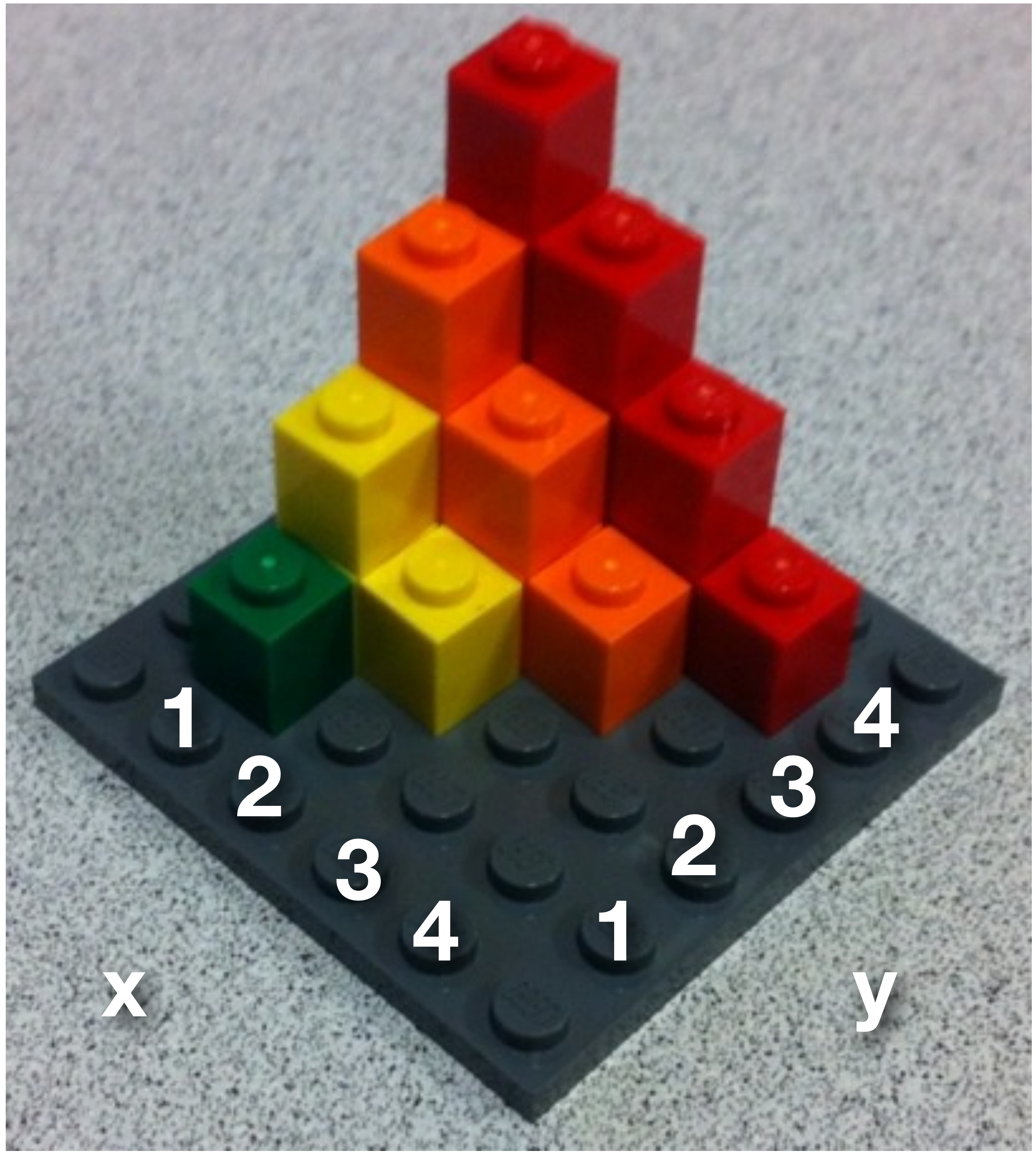
f(x)

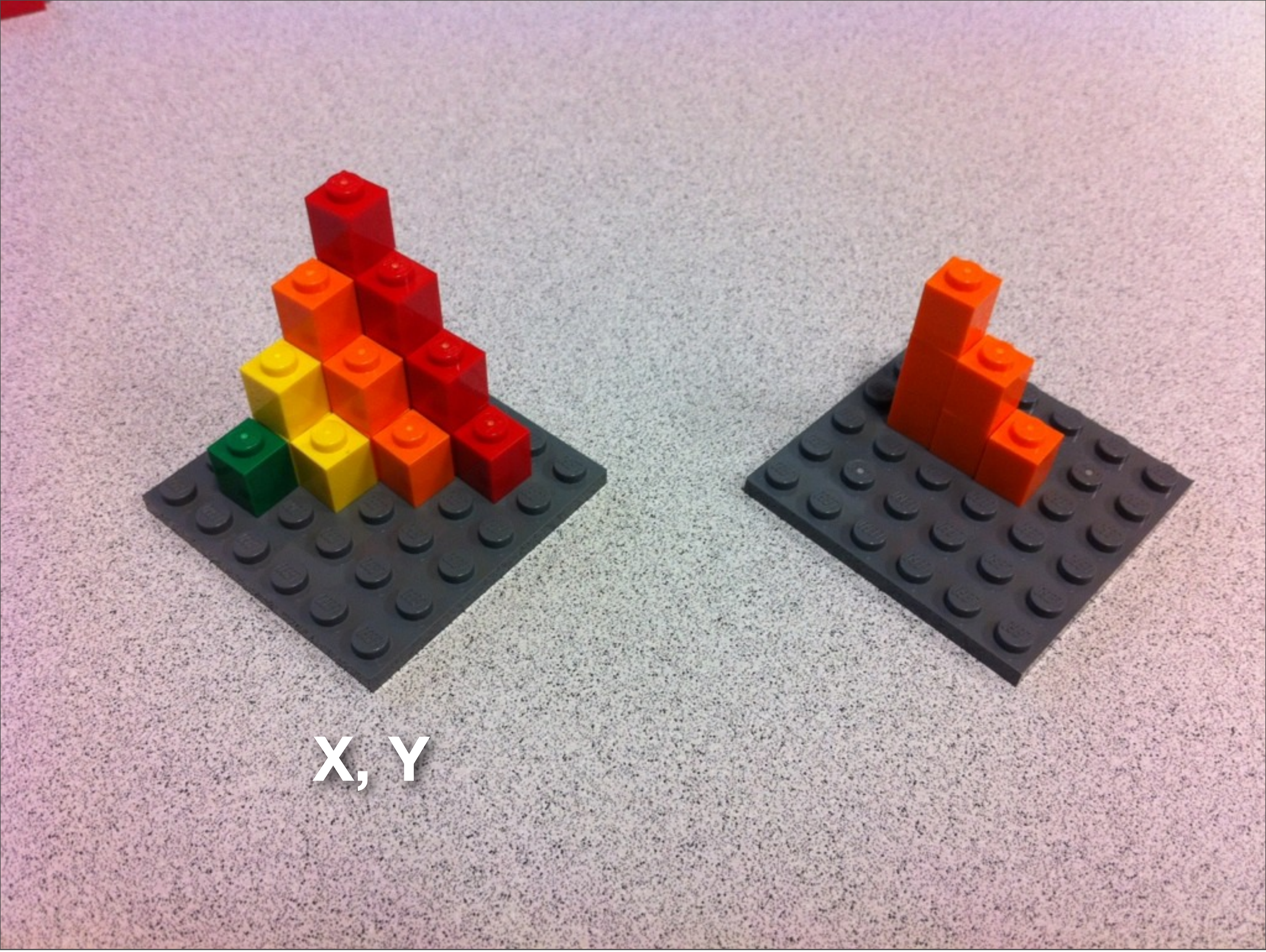


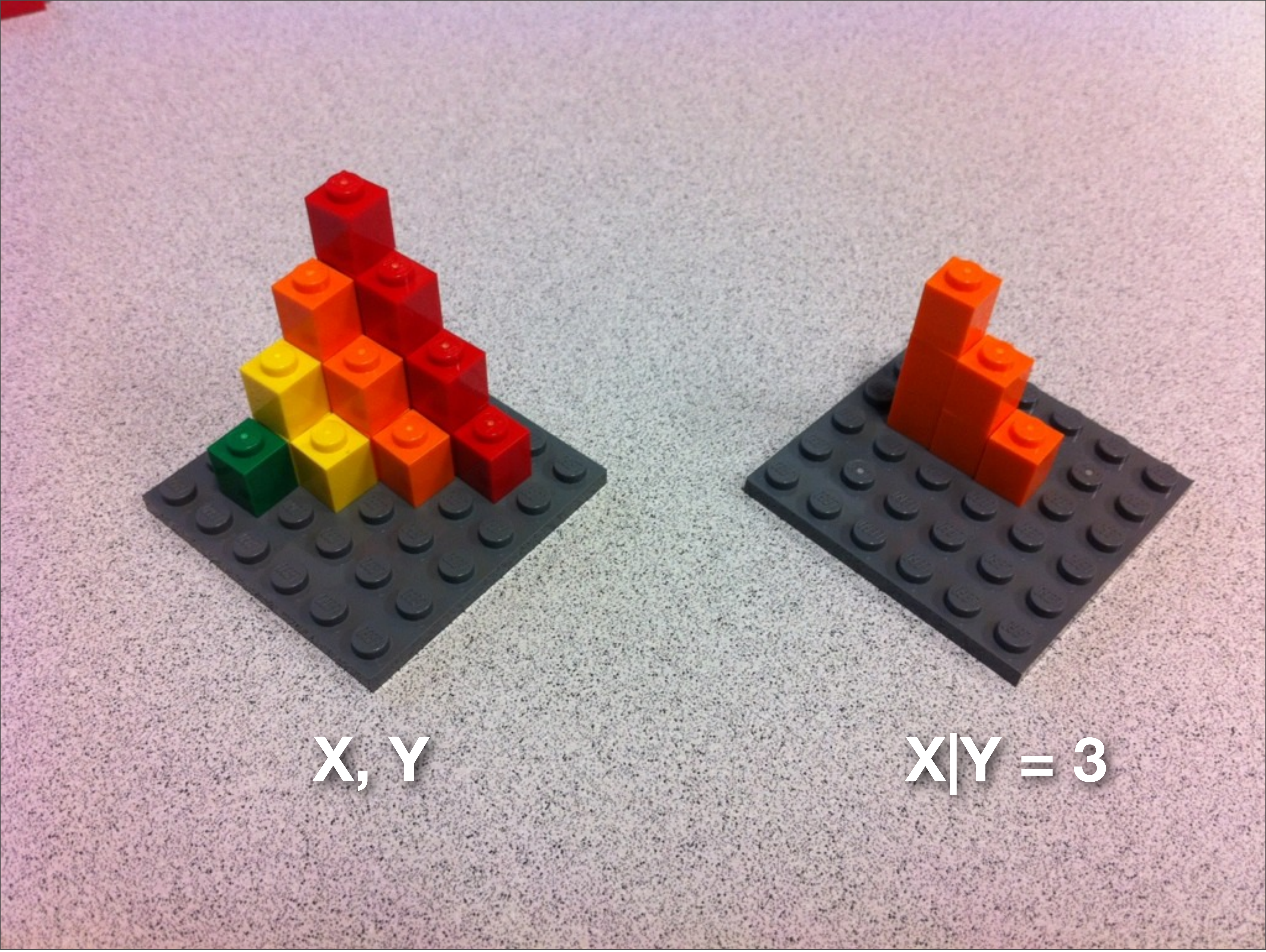
f(y)

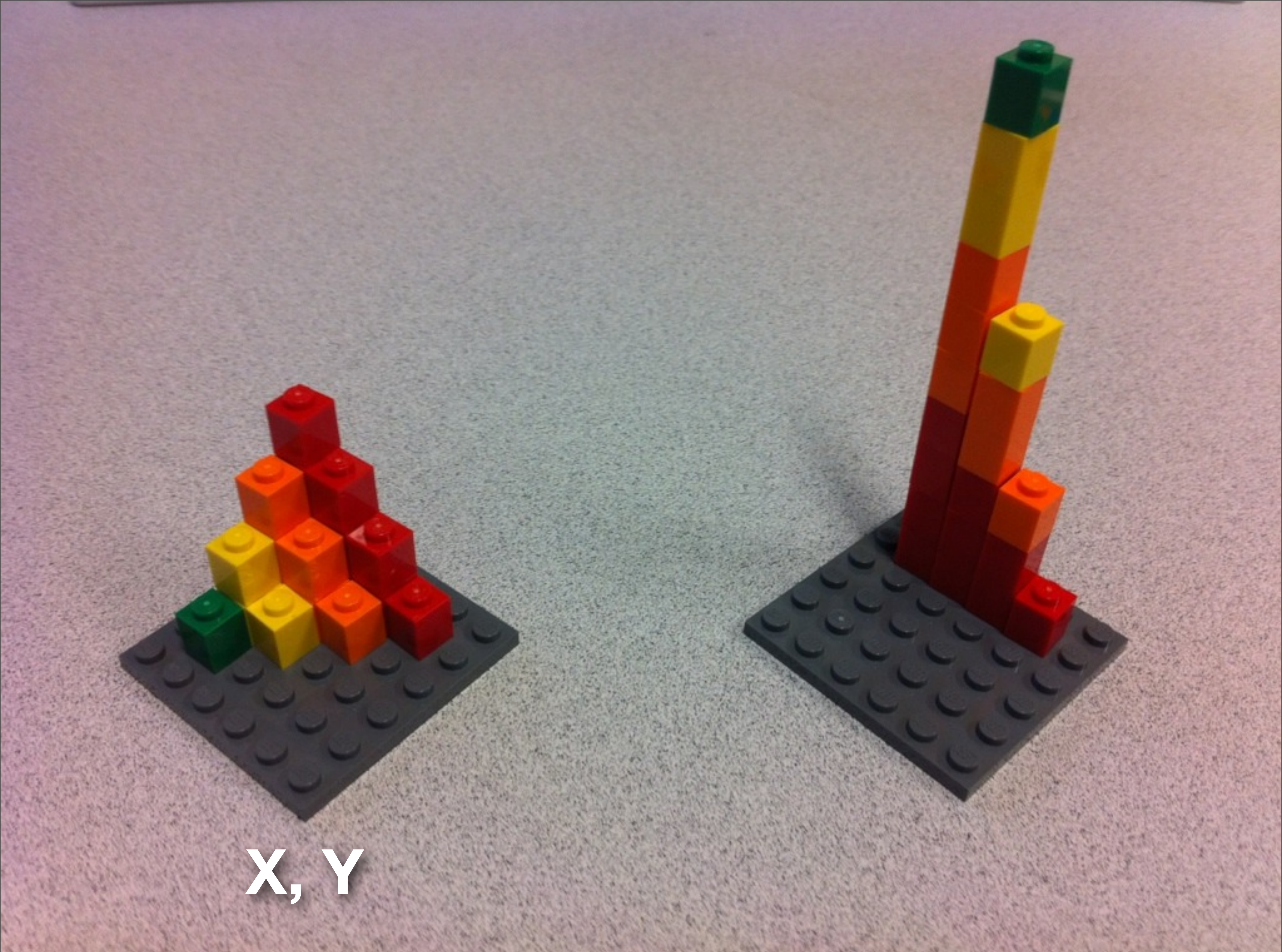


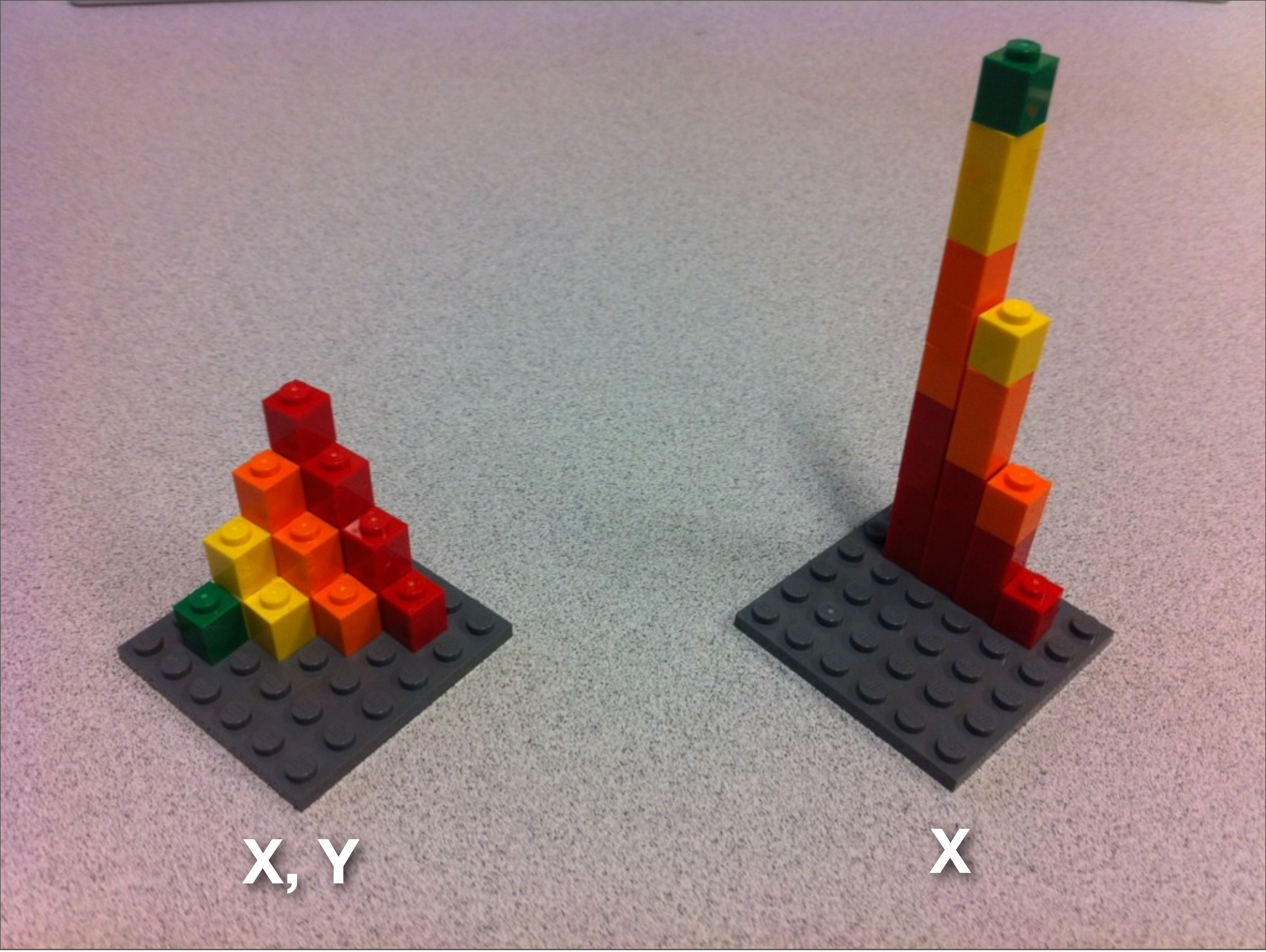

X, Y

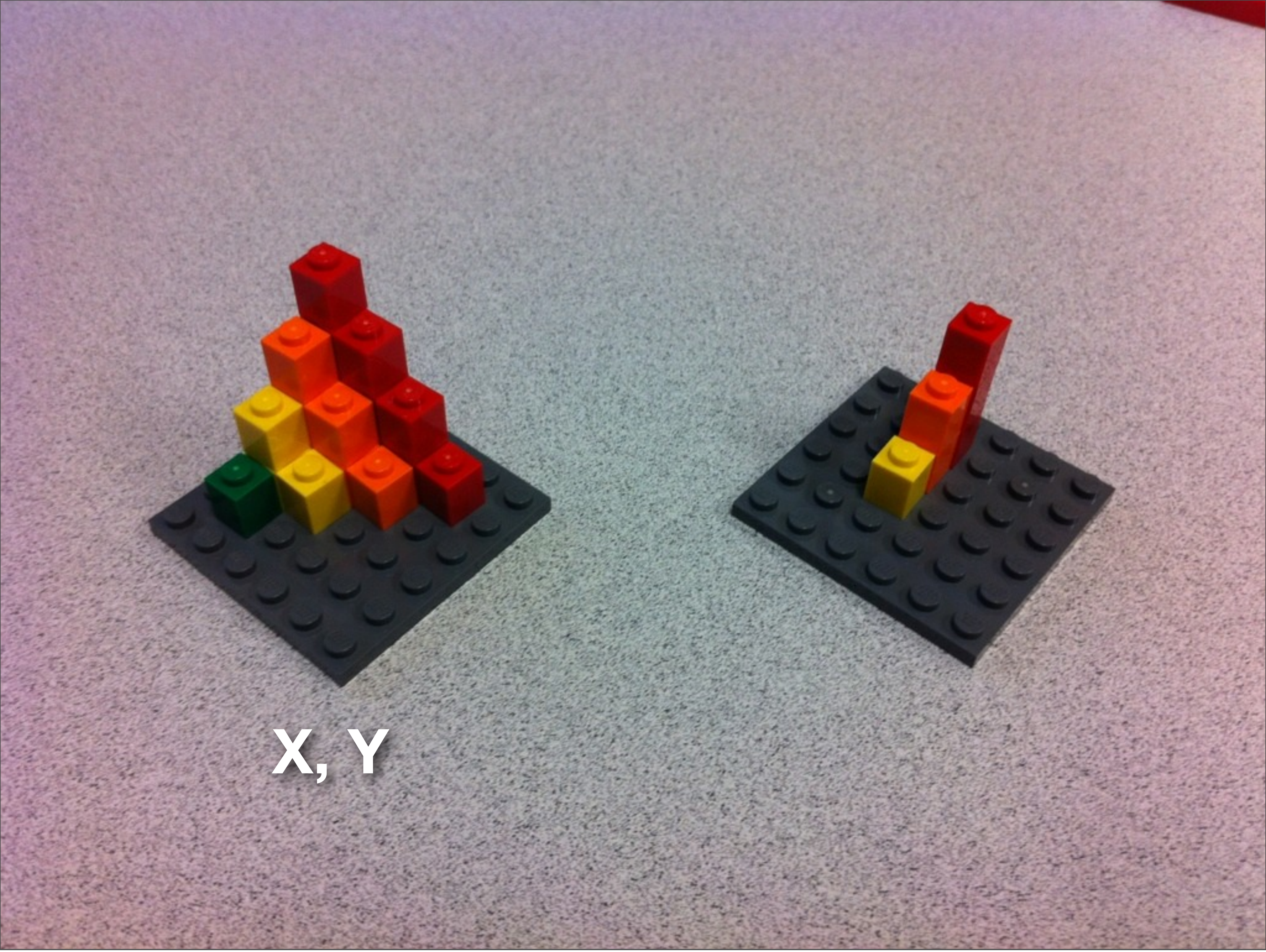

1
2
3
4
x
1
2
3
4
y

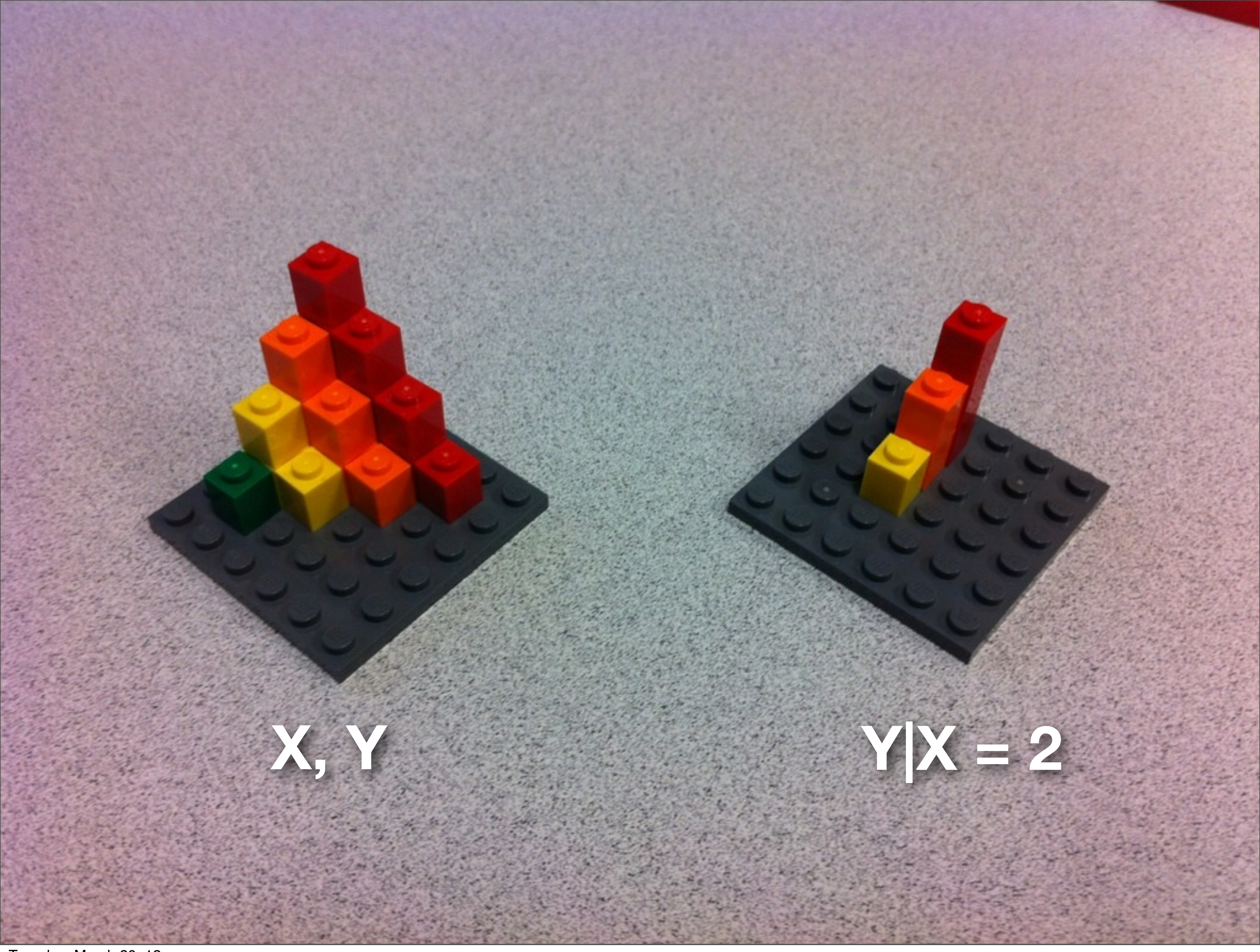

X, Y

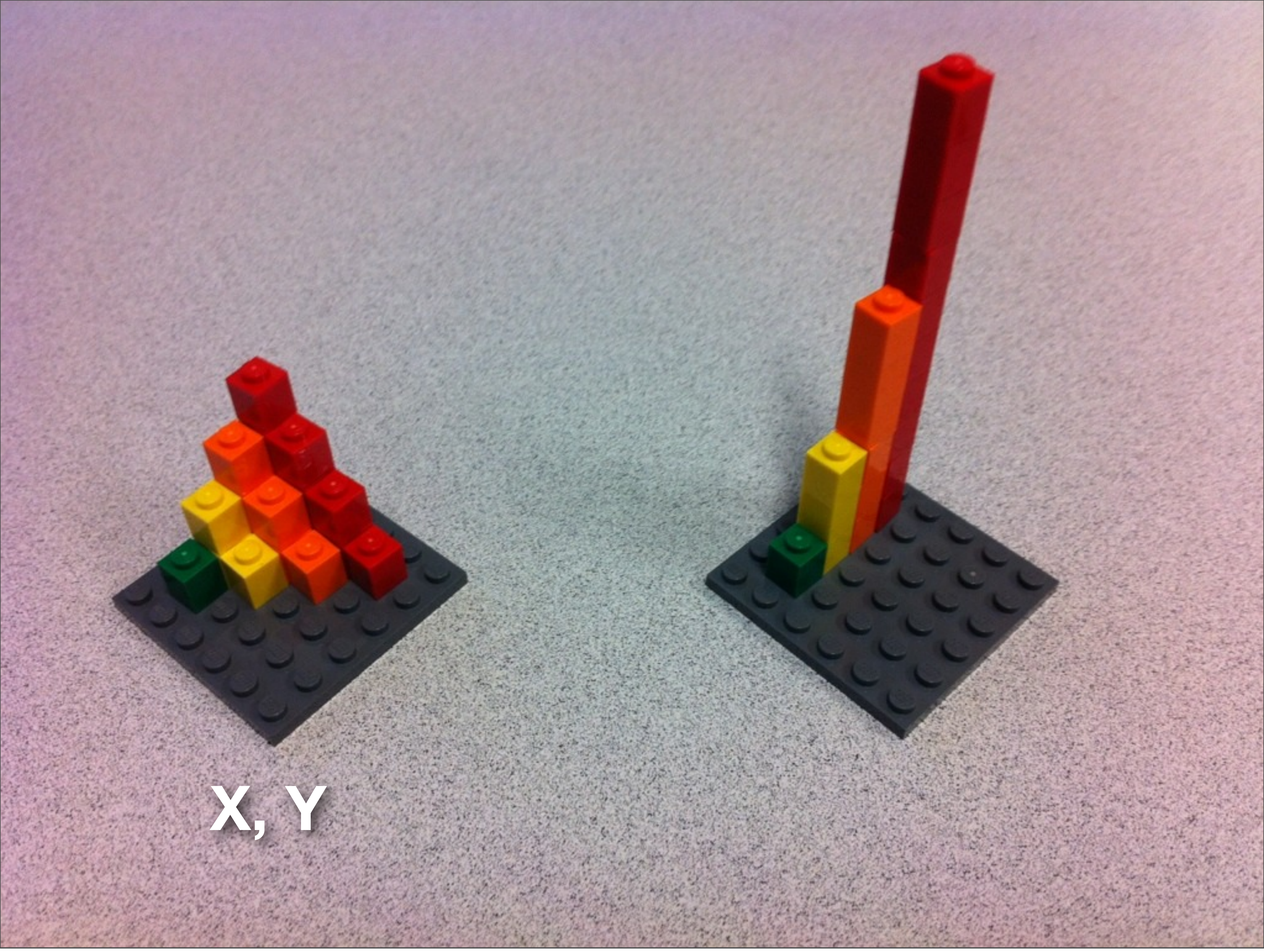

X, Y
X|Y = 3

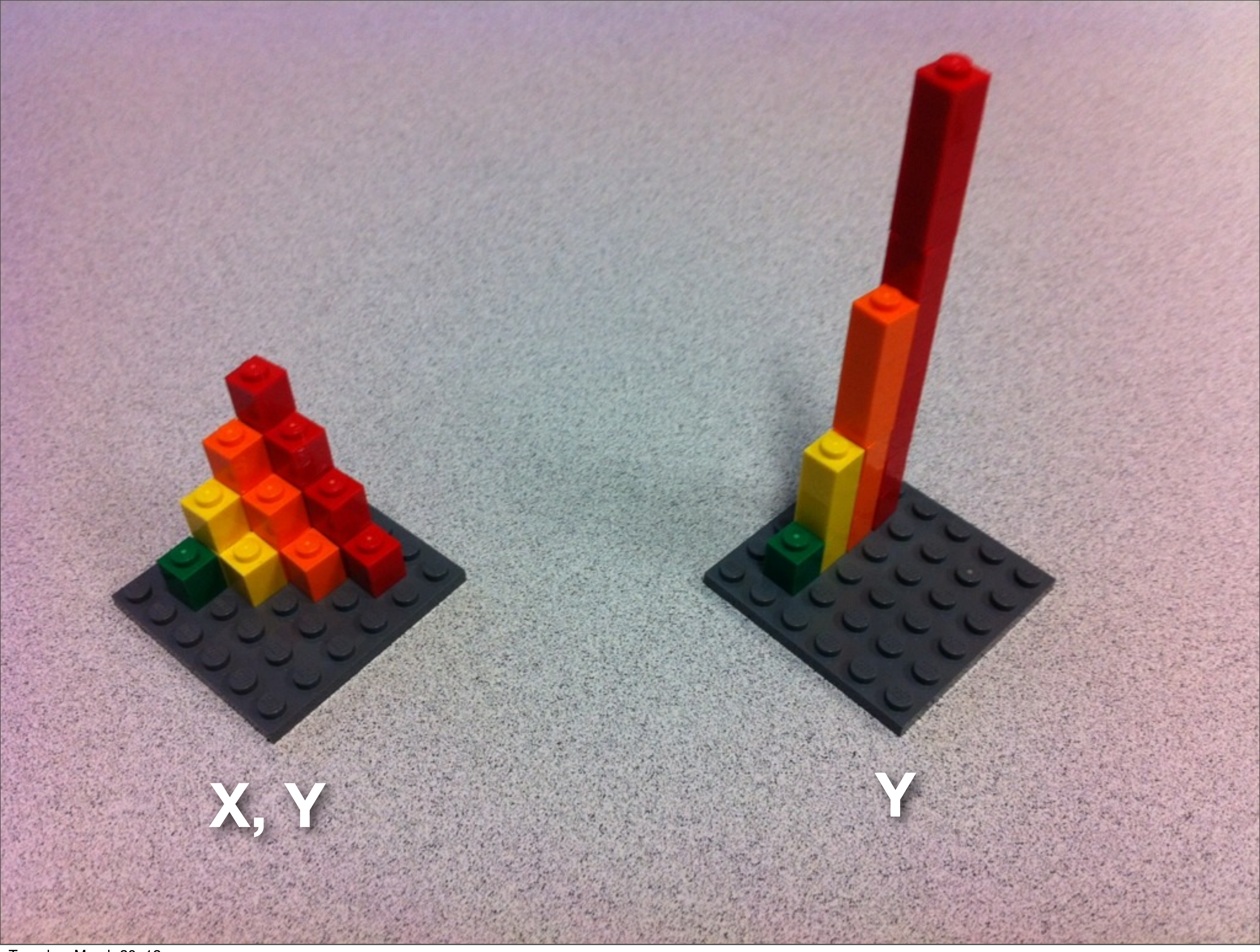

X, Y

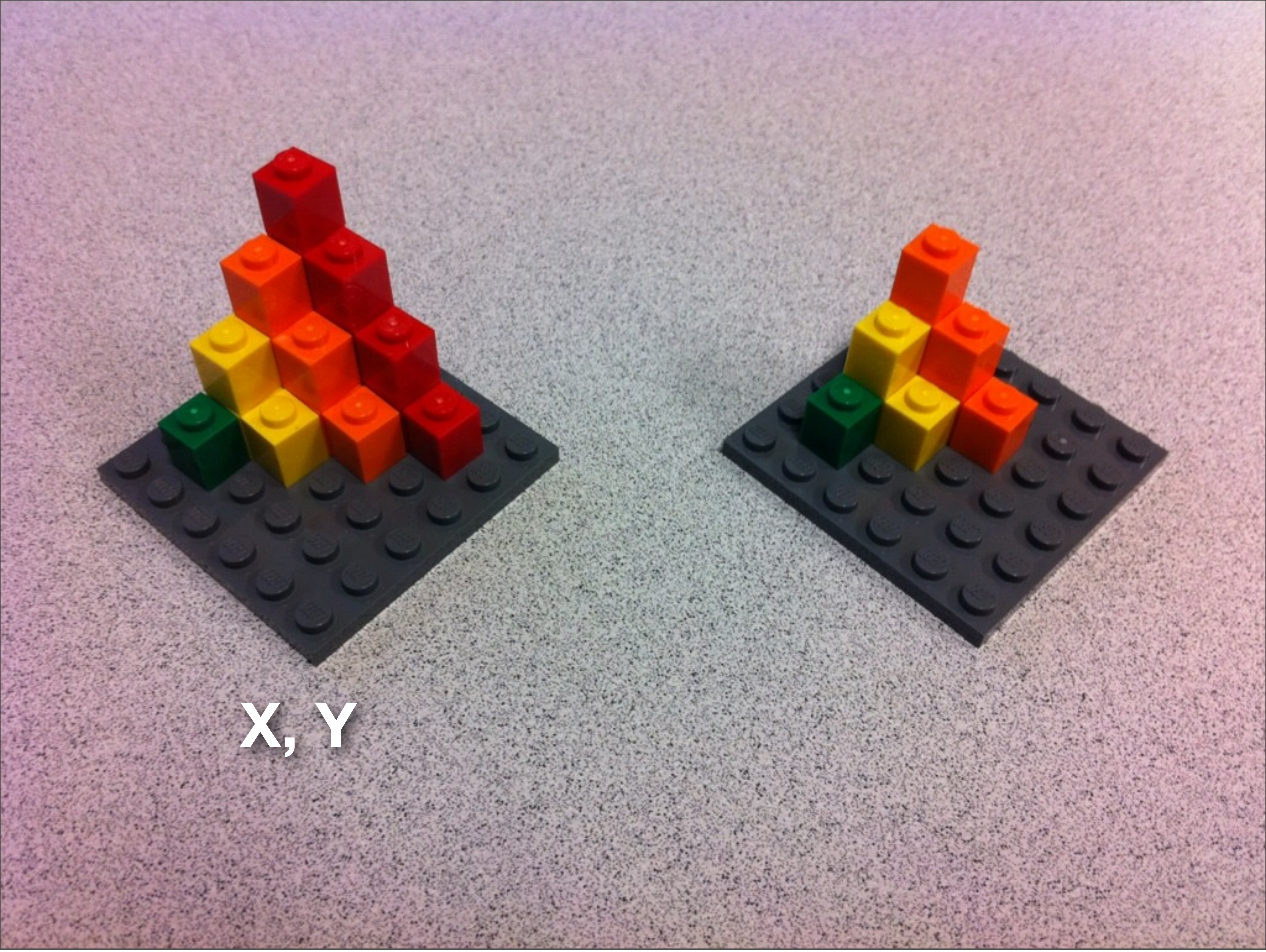

X, Y
X

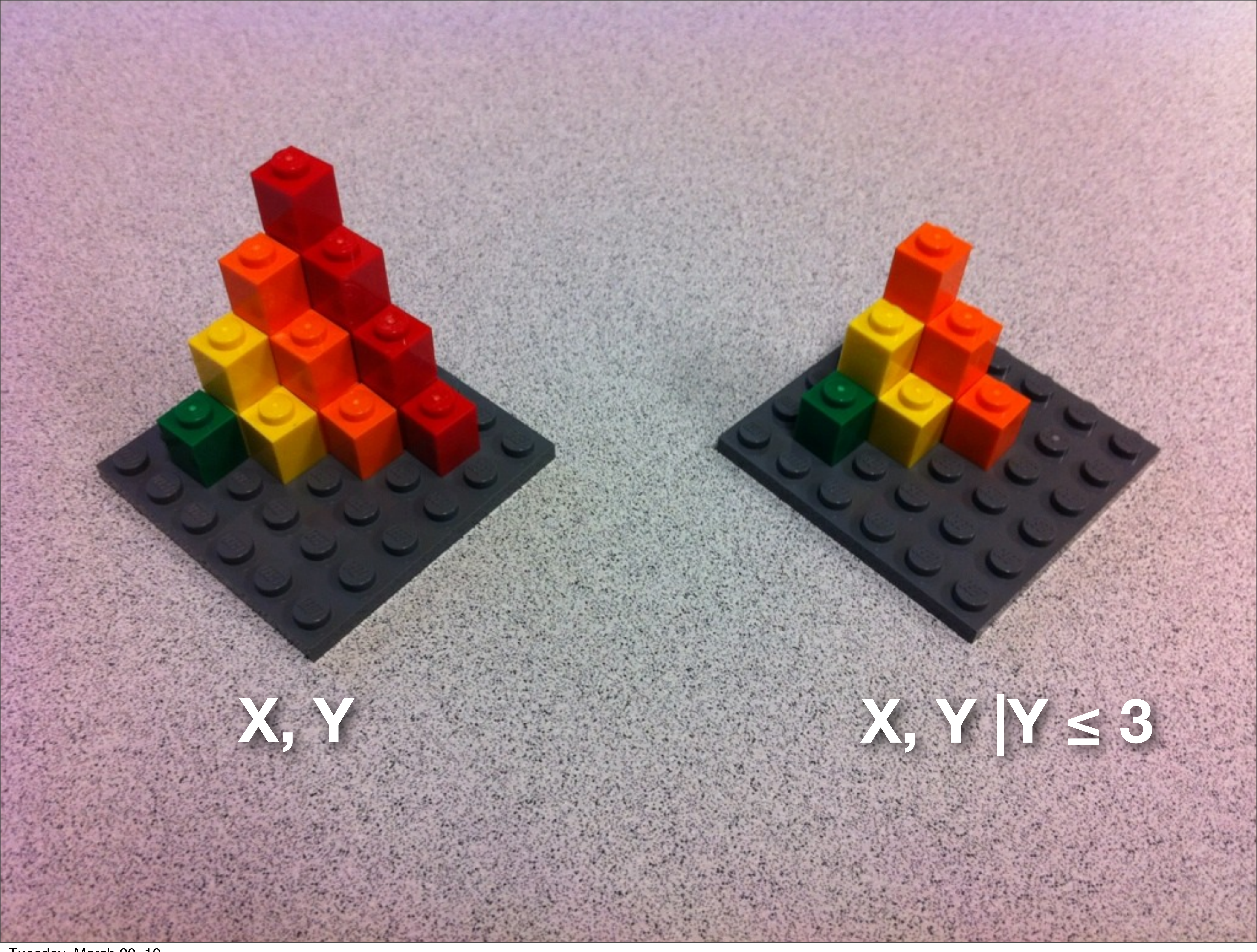

X, Y


X, Y
Y|X = 2


X, Y


X, Y
Y


X, Y


X, Y
X, Y |Y ≤ 3

2d 1d

Marginal

integration

Joint

fixing value

Conditional

$$f(x,y) = f(x|y)f(y)$$

$$f(x,y) = f(y|x)f(x)$$

X and Y are independent iff

$$f(x,y) = f(x)f(y)$$

Your turn

Let $X \sim$ Exponential(60)

Let $Y \mid X = x \sim$ Bernoulli($p = 1 / x$)

What would $f(x, y)$ look like?

What sort of problem is this modelling?

What is $E(Y)$?

$$f(x,y) \geq 0 \quad \forall\,(x,y) \in S$$

$$\int_{-\infty}^{\infty} \int_{\infty}^{\infty} f(x,y)\ dy\ dx = 1$$

$$P((X,Y) \in A) = \iint_A f(x,y)\,dx\,dy$$

$$P(x_1 < X < x_2, y_1 < Y < y_2) =$$

$$\int_{x_1}^{x_2} \int_{y_1}^{y_2} f(x,y)\,dy\,dx$$

$$E(u(X,Y)) = \iint_S u(x,y)f(x,y)\,dx\,dy$$

$$E(X) = \iint_S xf(x,y)\,dx\,dy$$

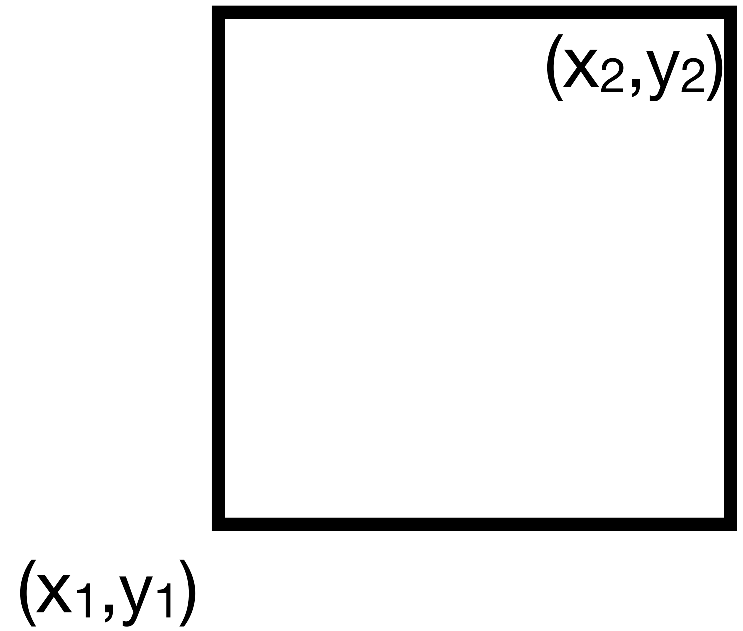
$$E(Y) = \iint_S yf(x,y)\,dx\,dy$$

Your turn

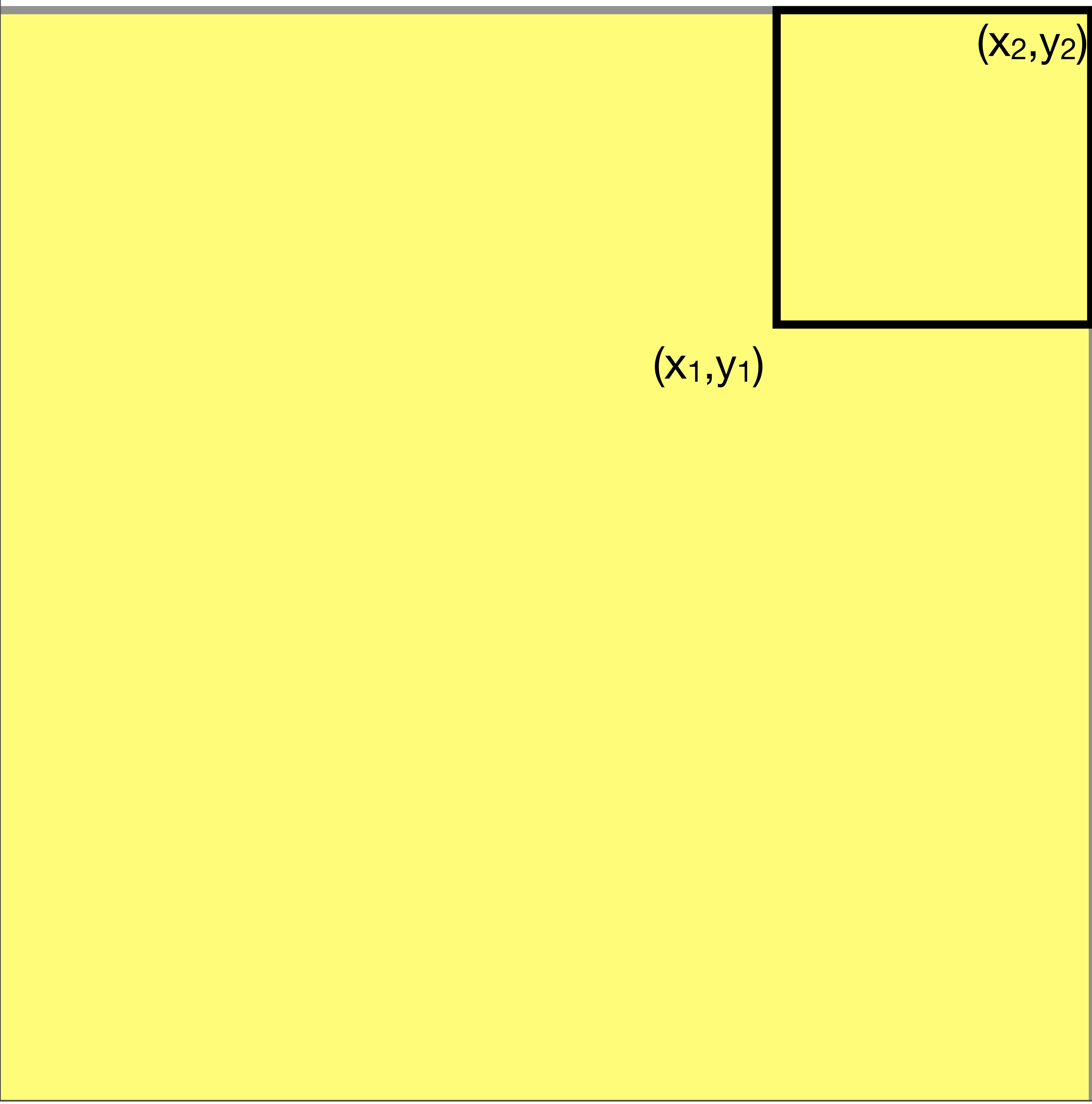
What do you think expectation would look like for a bivariate discrete random variable?

What about a bivariate distribution where one margin was discrete and one was continuous?

$P(x_1 < X < x_2, y_1 < Y < y_2) =$

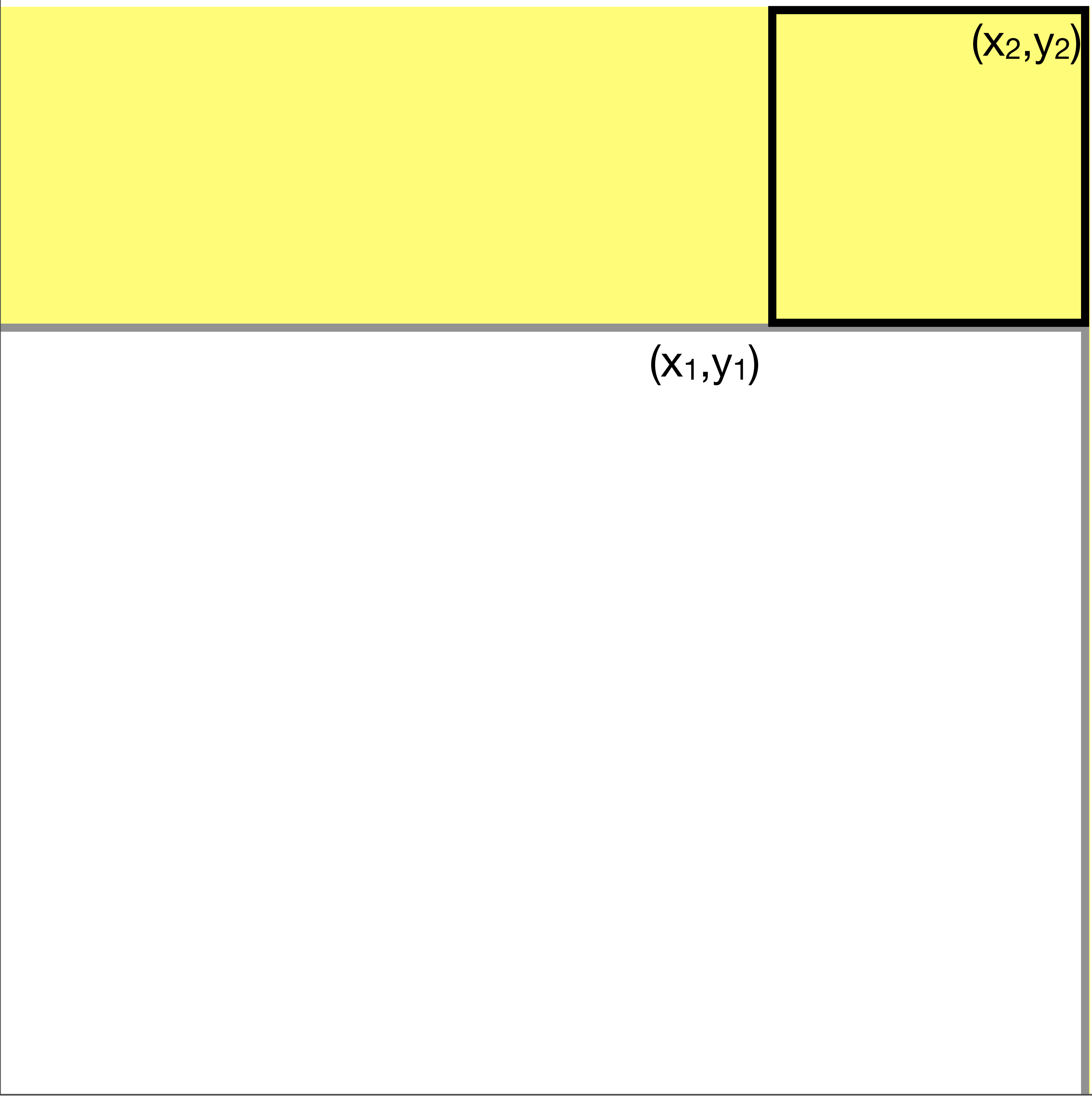


$$P(x_1 < X < x_2,\ y_1 < Y < y_2) =$$



$$F(x_2,\ y_2)$$

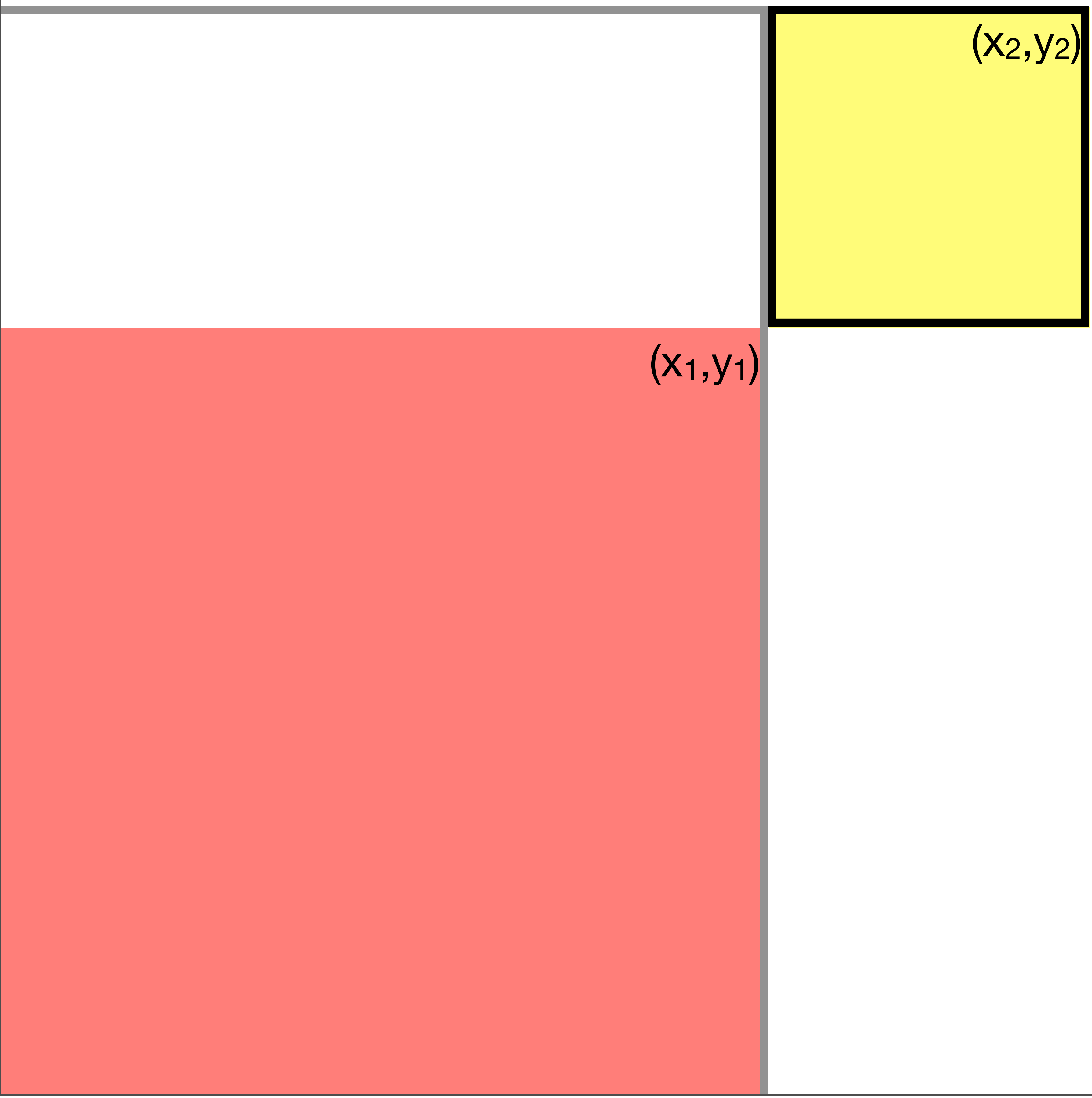
$P(x_1 < X < x_2, y_1 < Y < y_2) =$



$F(x_2, y_2)$

$\mathbf{- F(x_1, y_2)}$

$$P(x_1 < X < x_2, y_1 < Y < y_2) =$$

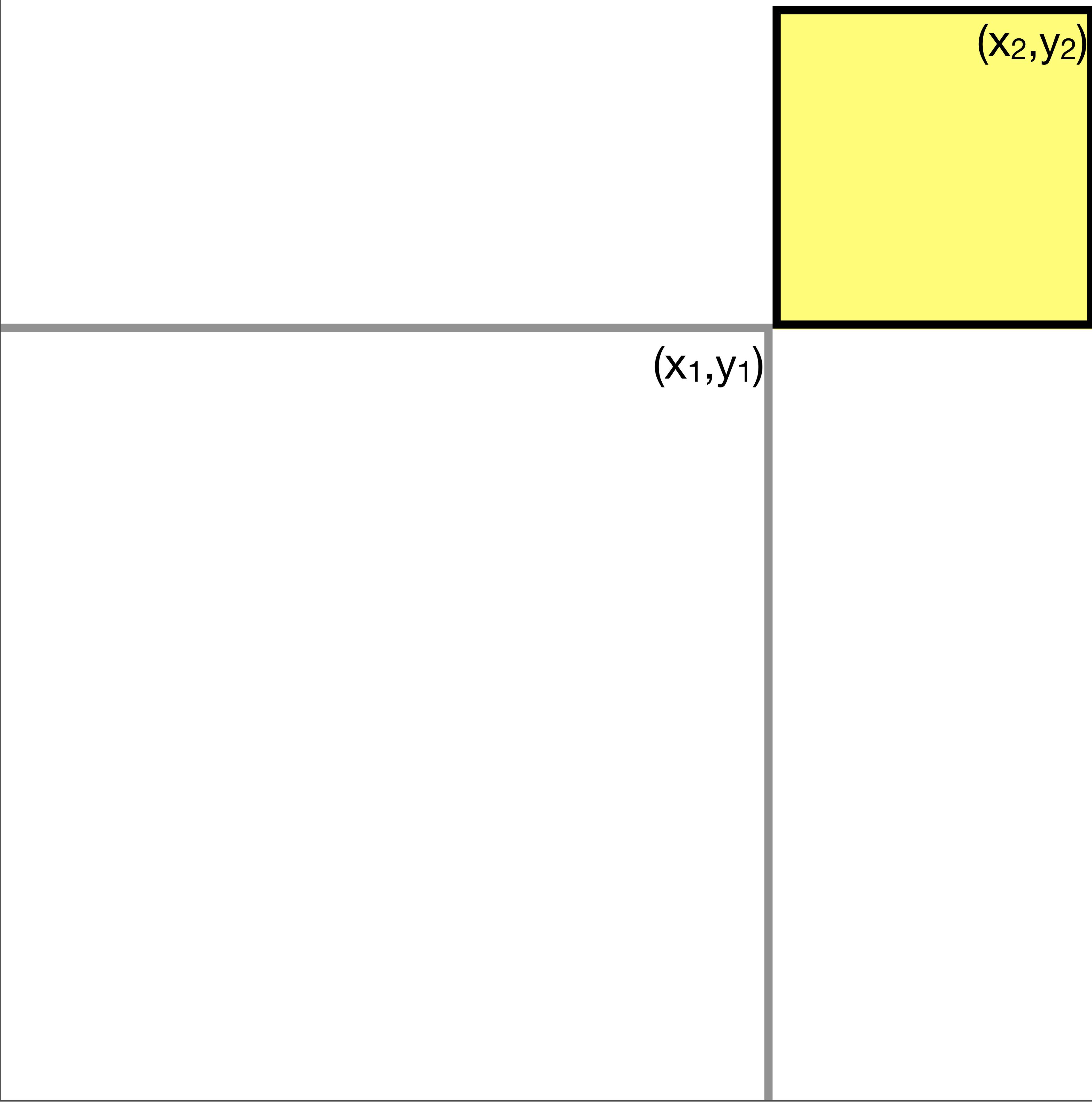


$$F(x_2, y_2)$$

$$- F(x_1, y_2)$$

$$\mathbf{- F(x_2, y_1)}$$

$$P(x_1 < X < x_2,\ y_1 < Y < y_2) =$$



$$F(x_2, y_2)$$
$$- F(x_1, y_2)$$
$$- F(x_2, y_1)$$
$$+ \mathbf{F(x_1, y_1)}$$

Transformations

	1d	2d
Change of variables	Must have inverse. Procedure easy	Must have inverse. Inverses can be tricky
Distribution function technique	Always works. Use definition of cdf	Easy, when it works

Also remember relationship between uniform distribution and any univariate distribution

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

Sequences

Important terms

iid

LLN

Chebyshev

CLT

Important tools

Limits

MGFs

Feedback