# Stat310

## Transformations

Hadley Wickham

1. Homework
2. Normal distribution and tables
3. Transformations

# Homework

Apologies for the error in the homework.

If I do make a correction, you can always choose to answer either version.

**New policy:** First person to report an error in the homework before 9am Monday will get bonus points.

# Normal distribution

# Parameters and pdf

The normal distribution has two parameters:

$\mu \in R$, the mean

$\sigma > 0$, the standard deviation
(or $\sigma^2 > 0$, the variance)

And has pdf...

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

Is this a valid pdf?

Note that this pdf is symmetric about mean!

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

Is this a valid pdf?

```
integrate 1/(sigma sqrt(2 pi)) e ^ (-(x- mu)^2 /
(2(sigma^2))) dx from -infinity to infinity
```

Note that this pdf is symmetric about mean!

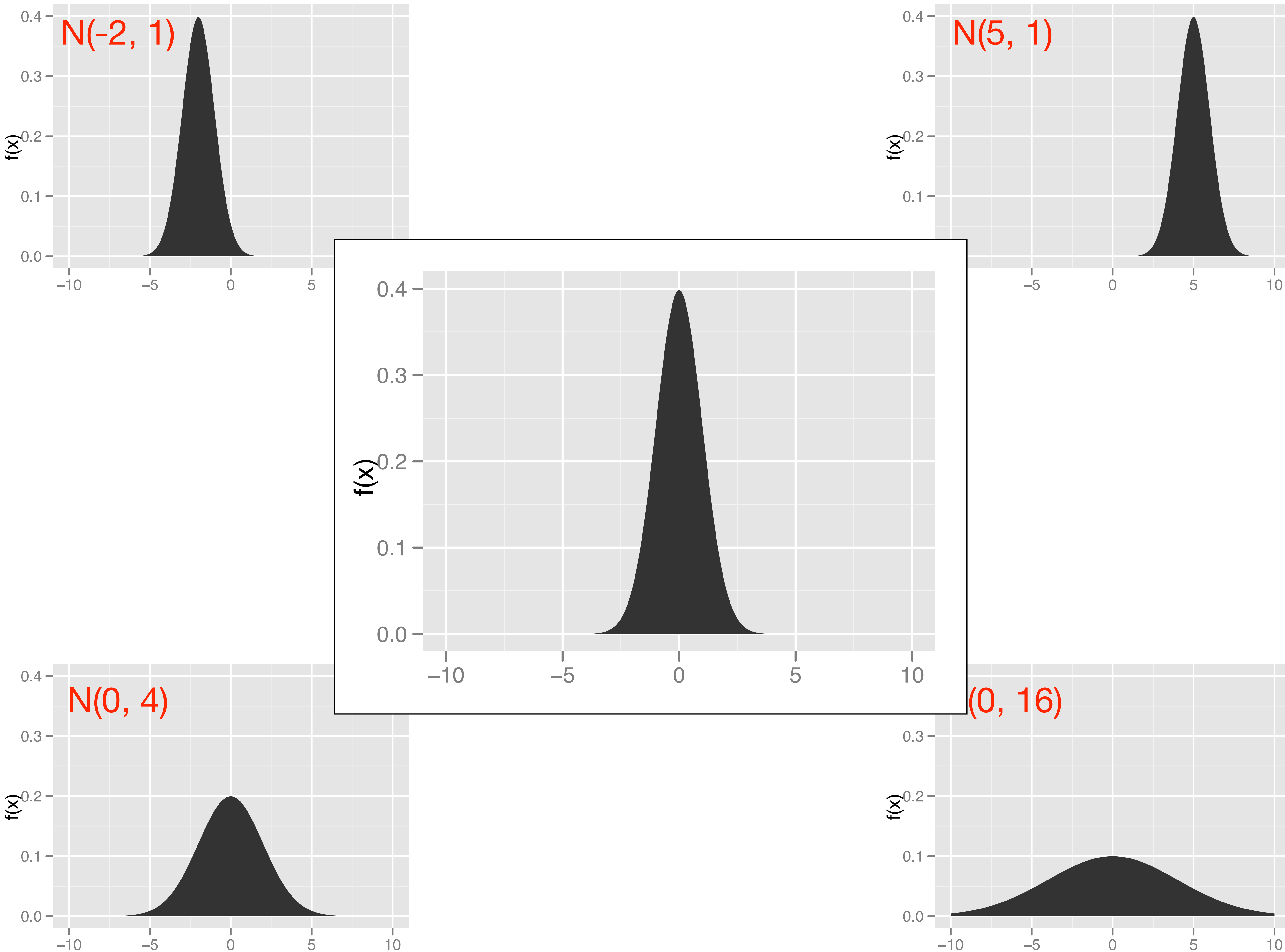
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Is this a valid pdf?

```
integrate 1/(sigma sqrt(2 pi)) e ^ (-(x- mu)^2 /
(2(sigma^2))) dx from -infinity to infinity
```

```
integrate 1/(sqrt(2 pi)) e ^ (-x^2 / 2) dx from -
infinity to infinity
```

Note that this pdf is symmetric about mean!


N(-2, 1)
N(5, 1)
N(0, 4)
(0, 16)
f(x)
0.0
0.1
0.2
0.3
0.4
−10
−5
0
5
10

# When to use

Empirically arises in many real situations, typically when a measurement can be thought of as the sum or mean of other measurements.

In a few weeks we’ll see why this true theoretically (the CLT)

# Moments

$$M(t) = \exp(\mu t + \frac{1}{2}\sigma^2 t^2)$$

What is the mean and variance of the normal distribution?

# With wolfram alpha

```
d/dt e^(mu*t + 1/2 sigma^2 t^2) at t = 0

d^2/dt^2 e^(mu*t + 1/2 sigma^2 t^2) at t
= 0

d^2/dz^2 exp(mu*z + 1/2 sigma^2 z^2) at
z = 0
```

# Transformations

If $X \sim \text{Normal}(\mu, \sigma^2)$, and $Y = aX + b$

THEN $Y \sim \text{Normal}(a\mu + b, a^2\sigma^2)$

# Standard normal

If b = -μ and a = 1/σ, we often write

$Z = (X - \mu) / \sigma$

$Z \sim \text{Normal}(0, 1)$ = **standard normal**

We'll see why this is so important shortly

# Your turn

Assume the weight of a bag of m&m's is normally distributed with a mean of 50 grams with a variance of 4 grams.

What's the probability that you get a bag that weighs more than 55 grams? (How would you work it out?)

# Normal cdf

$$P(Z < z) = \Phi(z)$$

So commonly used that it has it’s own symbol.

But what is it?

```
integrate 1/(sigma sqrt(2 pi)) e ^ (-(t -
mu)^2 / (2(sigma^2))) dt from -infinity to
x
integrate 1/(sqrt(2 pi)) e ^ (-t^2 / 2) dt
from -infinity to x
```

# Tables

# Two approaches

If you don't have a closed form solution you can:

Look up the number in a table.

> Problem: you need a table for every combination of the parameters

Use a computer.

> Problem: you need a computer

# Standard normal

Fortunately, for the normal distribution, we can convert any random variable with a normal distribution to a **standard normal**.

This means we only need one table for any possible normal distribution. (For other distributions there will be multiple tables, and typically you will have to pick one with similar values to your example).

# Using the tables

Column + row = z

Find: $\Phi(2.94)$, $\Phi(-1)$, $\Phi(0.01)$, $\Phi(4)$

Tricks: $\Phi(0) = 0.5$, $\Phi(-x) = 1 - \Phi(x)$

Can also use in reverse: For what value of z is $P(Z < z) = 0.90$ ? i.e. What is $\Phi^{-1}(0.90)$?

Find: $\Phi^{-1}(0.1)$, $\Phi^{-1}(0.5)$, $\Phi^{-1}(0.65)$, $\Phi^{-1}(1)$

# Your turn

Assume the weight of a bag of m&m's is 50 grams with a variance of 4 grams.

What's the probability that you get a bag that weighs more than 55 grams?

(Convert to standard normal and then look up in table)

# Transformations

# Recall

| x | -1 | 0 | 1 | 2 | 3 |
|---|---|---|---|---|---|
| f(x) | 0.2 | 0.1 | 0.3 | 0.1 | 0.3 |

Let X be a discrete random variable with pmf f as defined above.

Write out the pmfs for:

$A = X + 2$    $B = 3X$    $C = X^2$  $D = 0 * X$

Remember: it’s the **sample space** that changes, not the probability

# Continuous

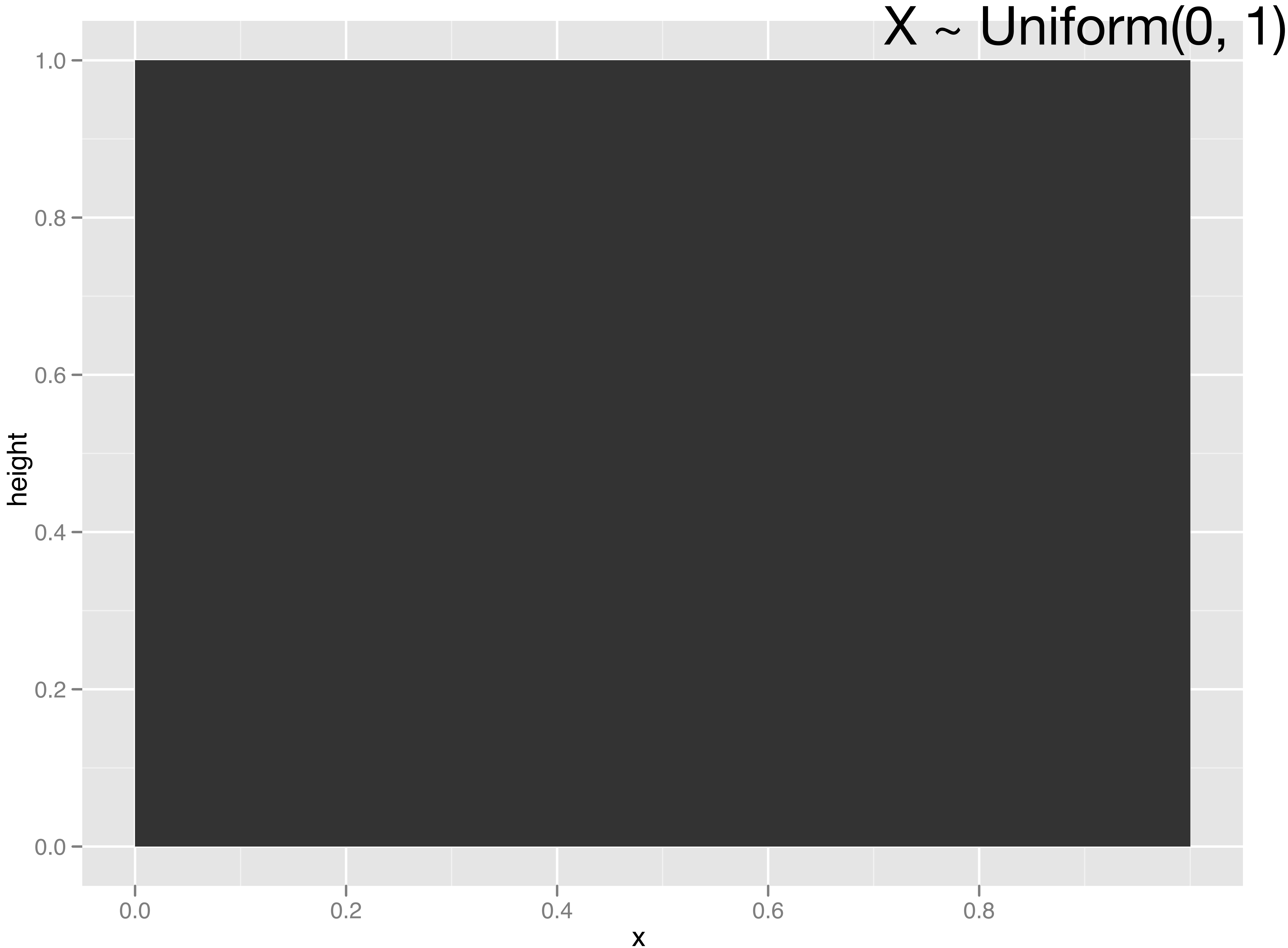
Let $X \sim \text{Unif}(0, 1)$

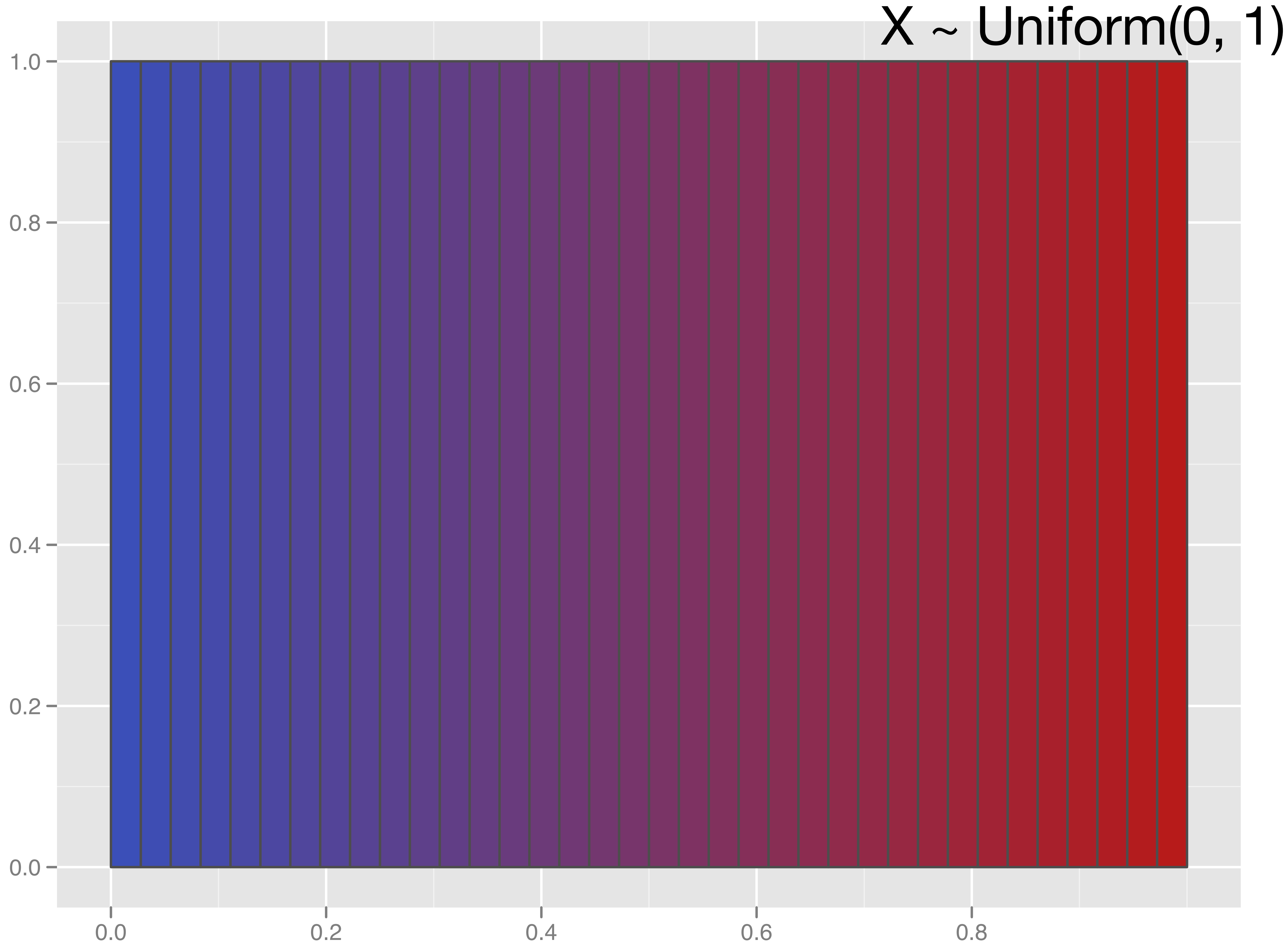
What are the distributions of the following variables?

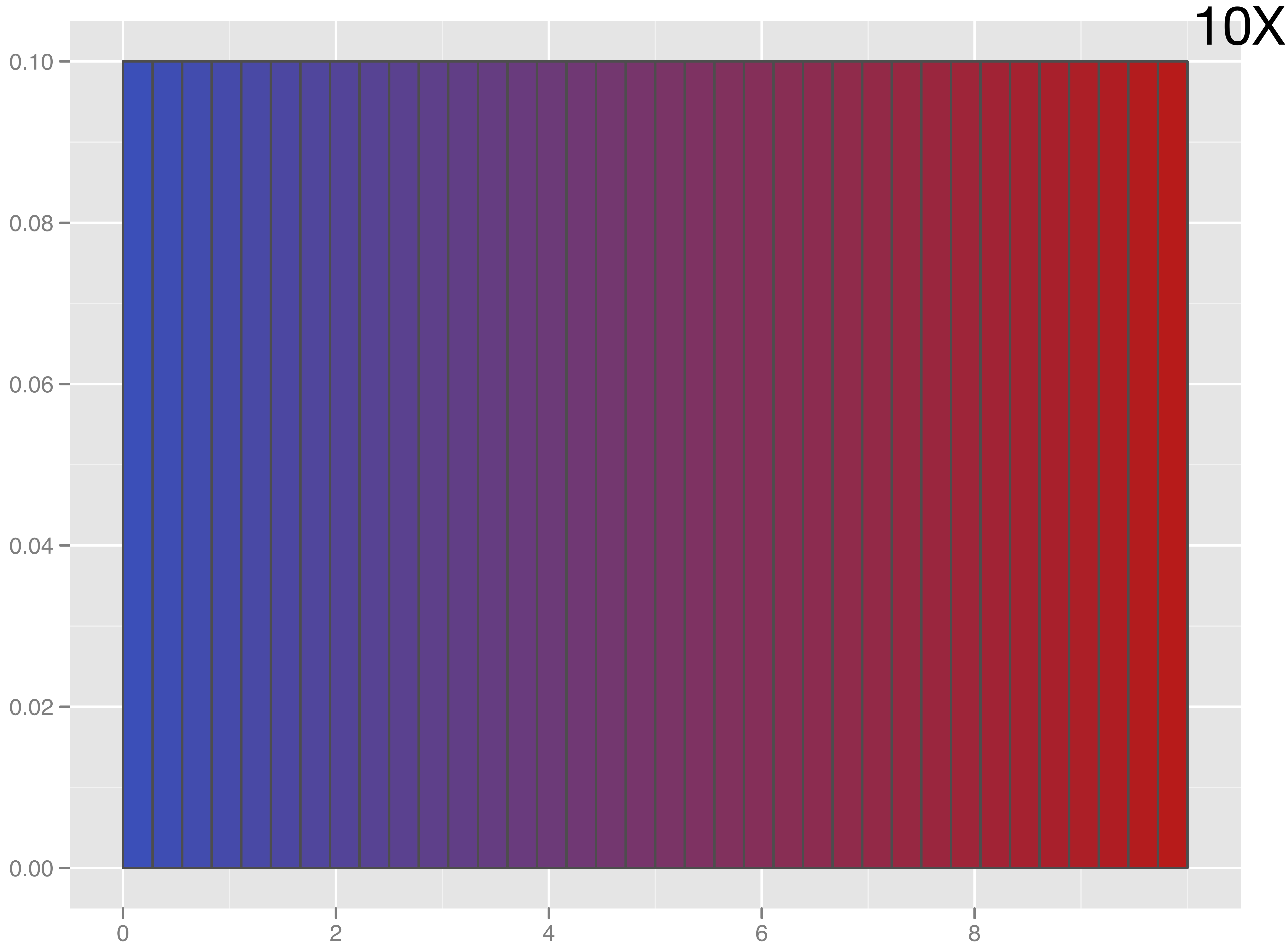
$A = 10X$

$B = 5X + 3$

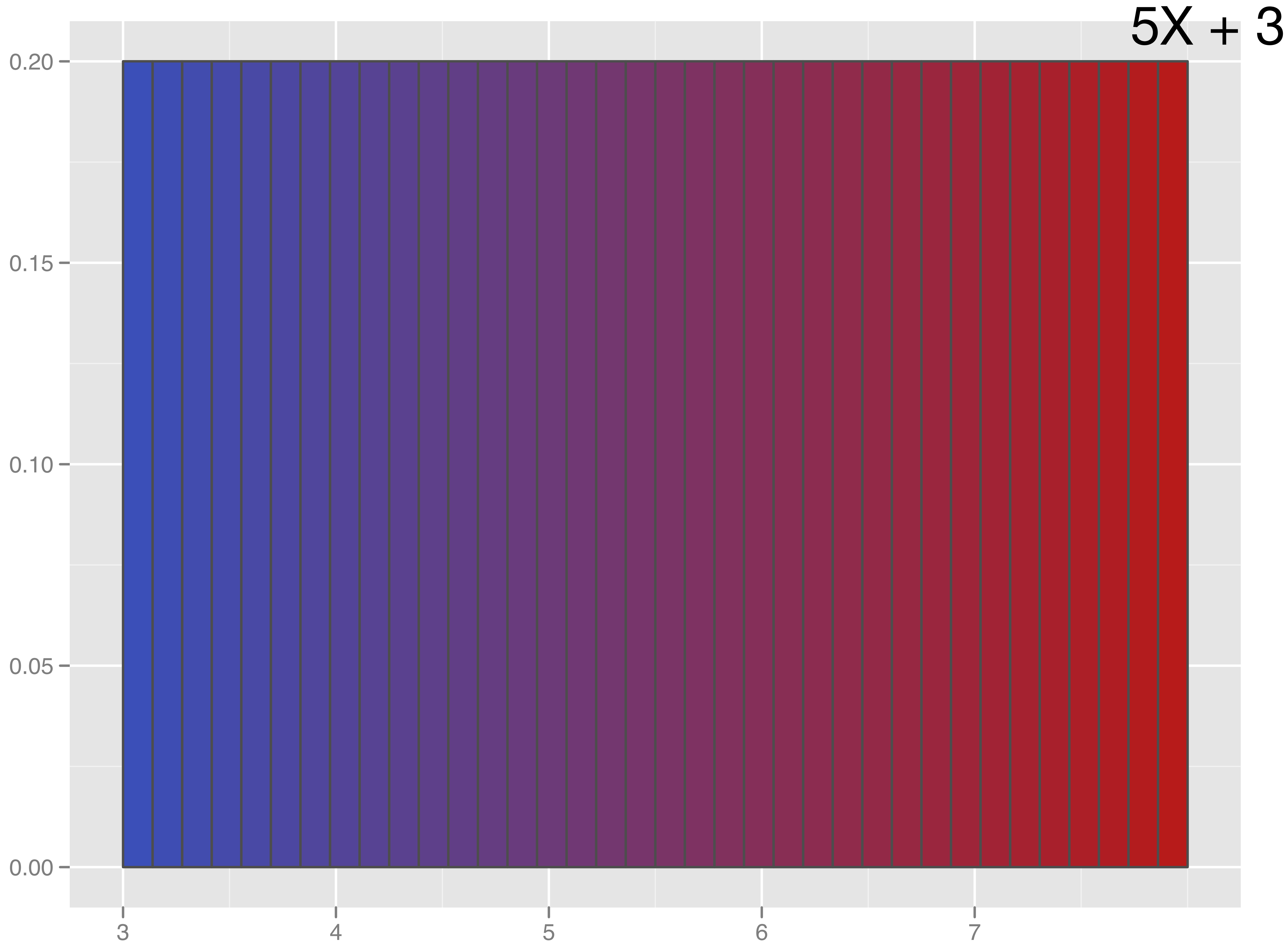
$C = X^2$

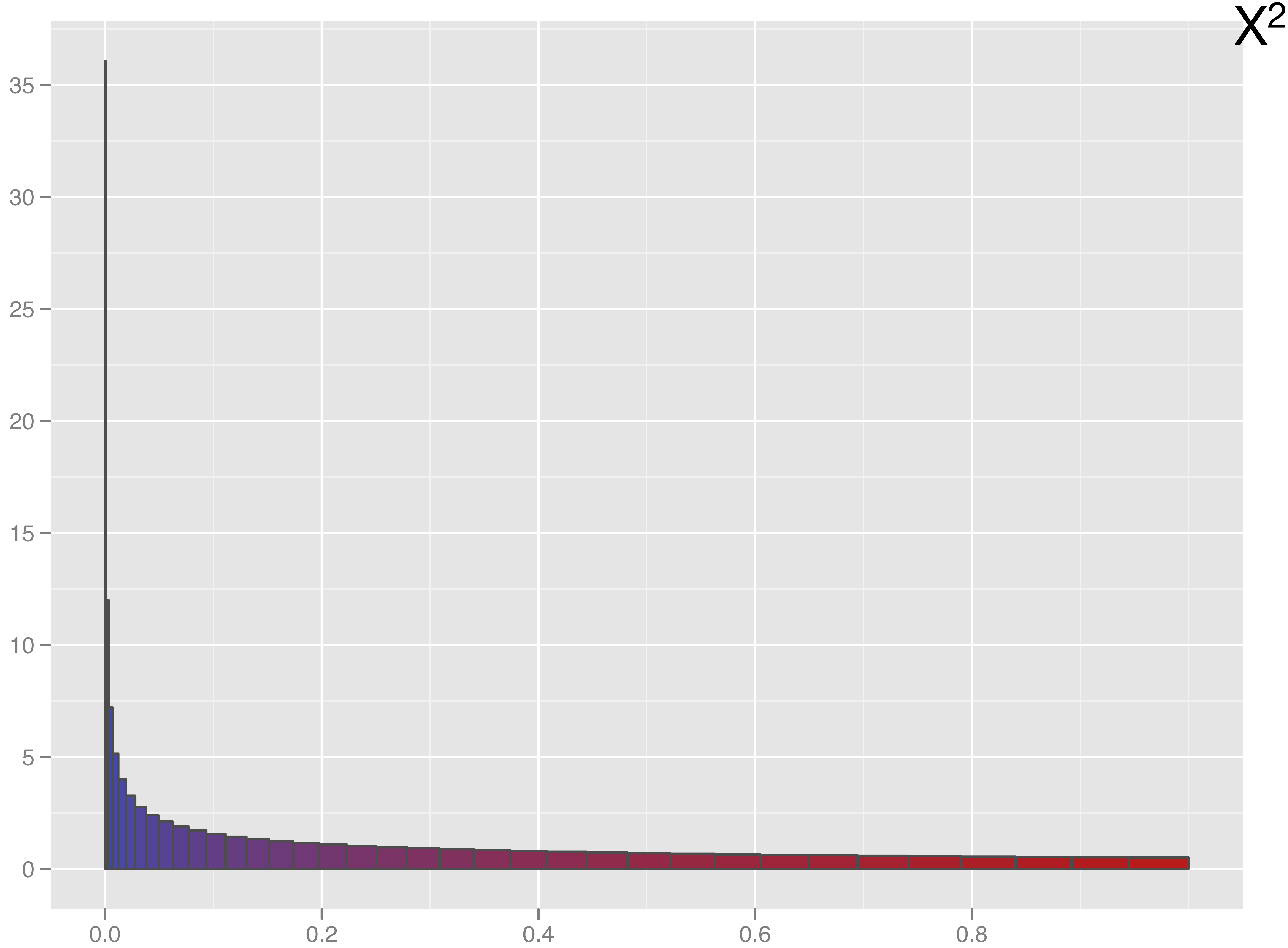

X ~ Uniform(0, 1)
1.0
0.8
0.6
0.4
0.2
0.0
height
0.0
0.2
0.4
0.6
0.8
x


X ~ Uniform(0, 1)
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8




Thursday, February 9, 12


5X + 3
0.20
0.15
0.10
0.05
0.00
3
4
5
6
7



X²
35
30
25
20
15
10
5
0
0.0
0.2
0.4
0.6
0.8

# Transformations

Distribution function technique

(always works, but hard)

Change of variable technique

(easy, but doesn't always work)

Much more useful for 2+ rvs

# Distribution function technique

# Distribution function technique

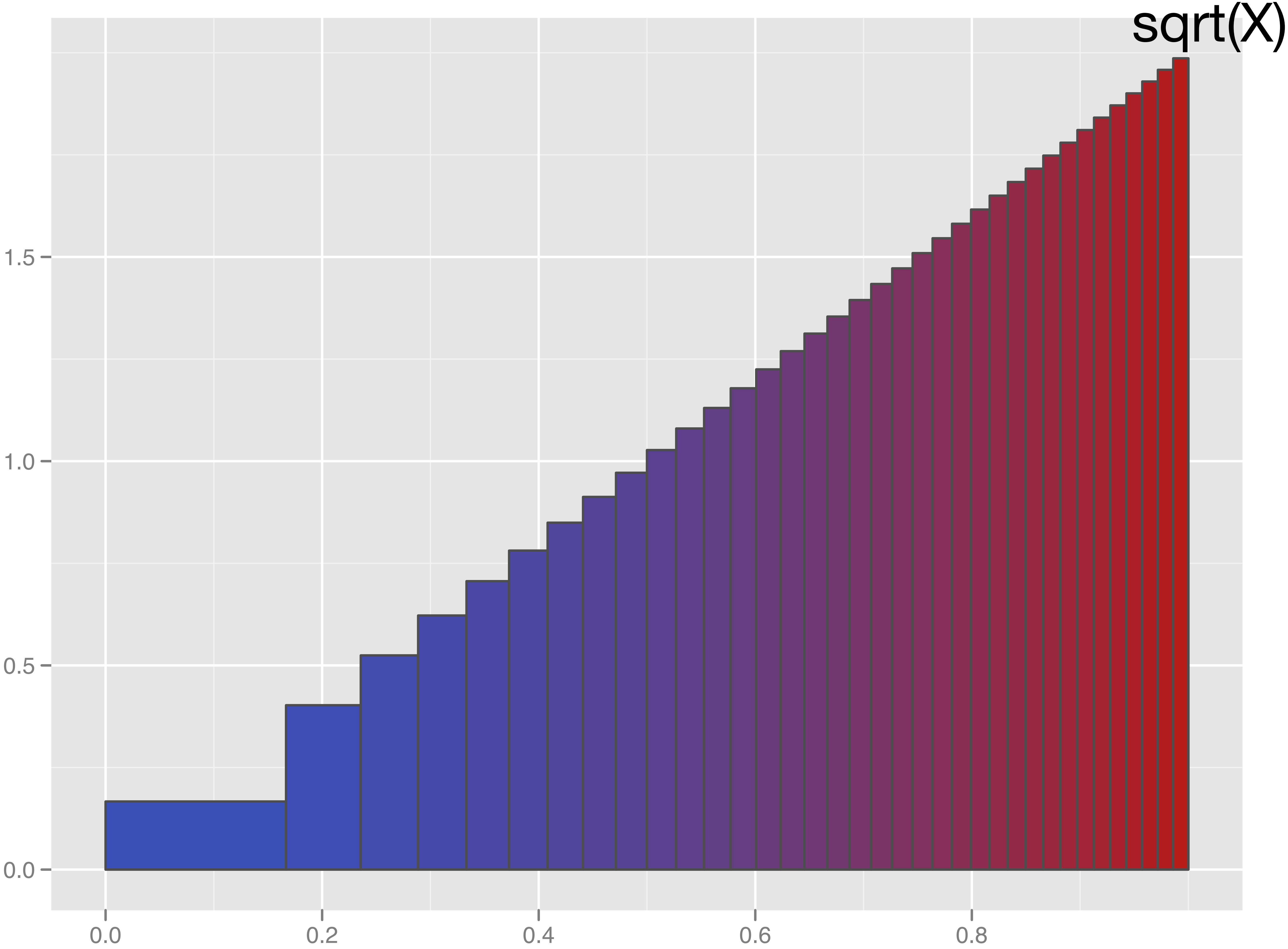
$Y = u(X)$

1. Figure out range
2. Write cdf of Y, then substitute for x
3. Find region in X such that $u(X) < y$
4. Integrate to compute cdf
5. Differentiate to get pdf

# Example

$X \sim \text{Unif}(0, 1)$

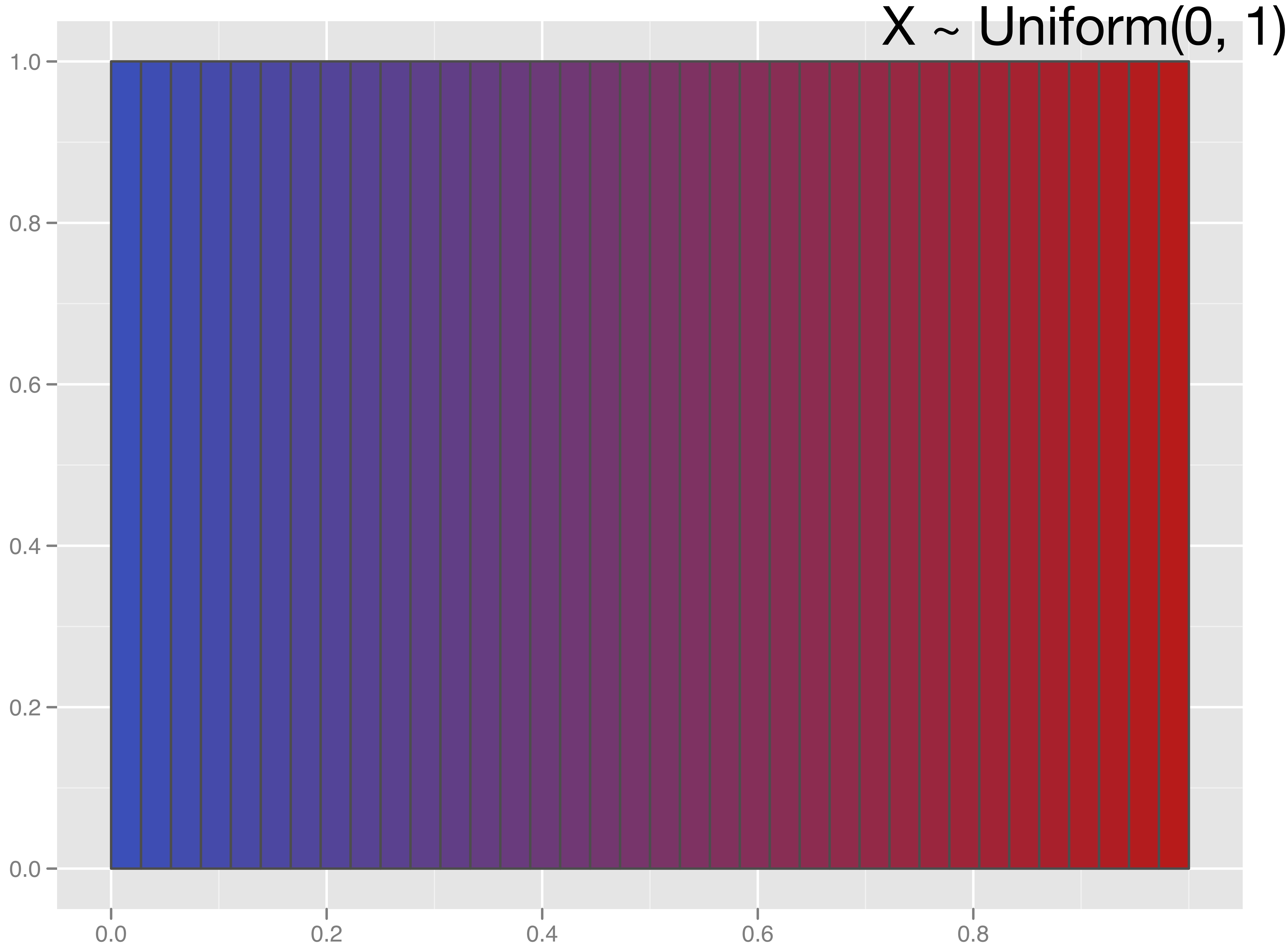
$Y = X^{0.5}$


sqrt(X)
1.5
1.0
0.5
0.0
0.0
0.2
0.4
0.6
0.8

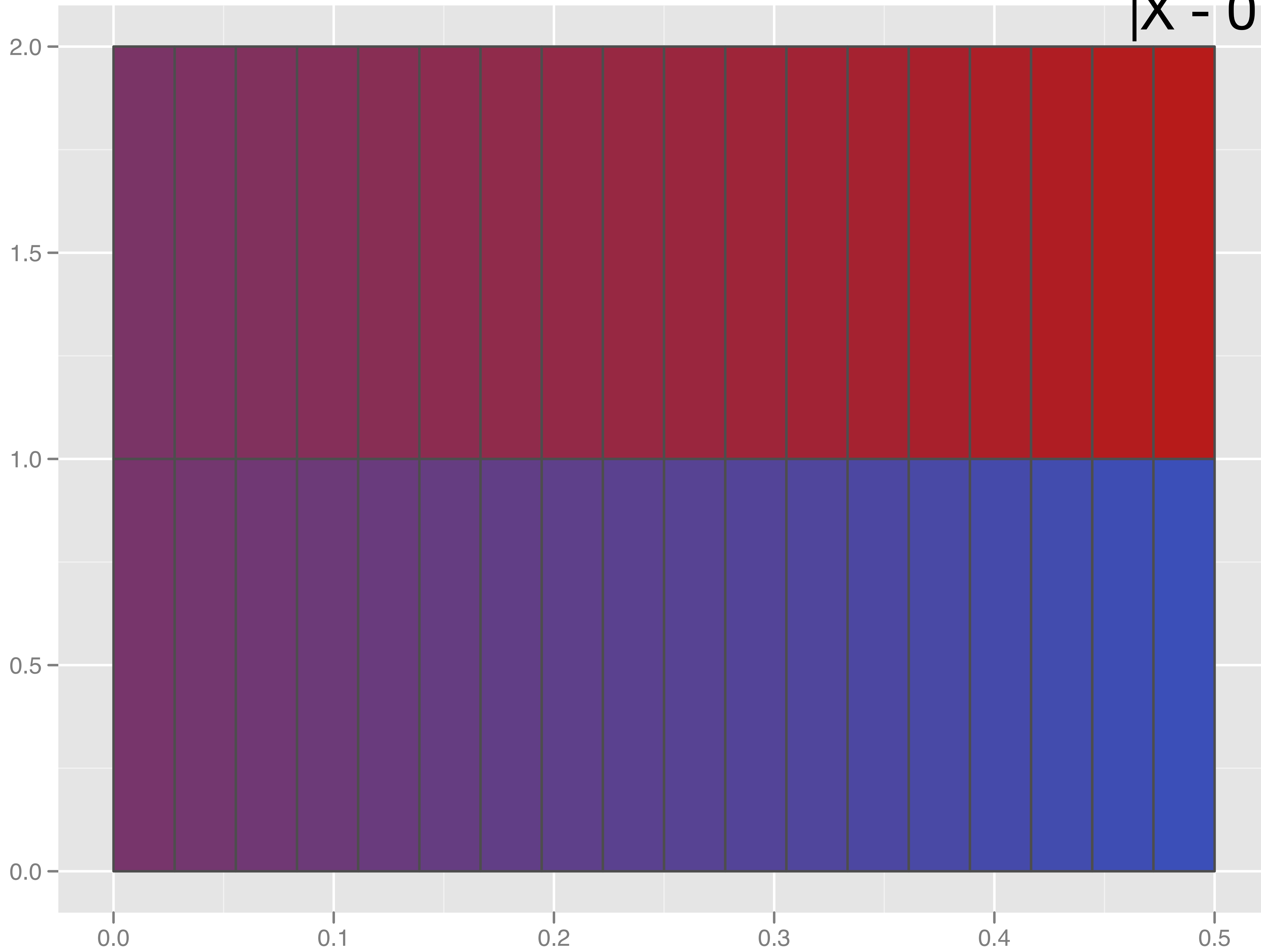
# Example

$X \sim \text{Unif}(0, 1)$

$Z = |X - 0.5|$


X ~ Uniform(0, 1)
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8

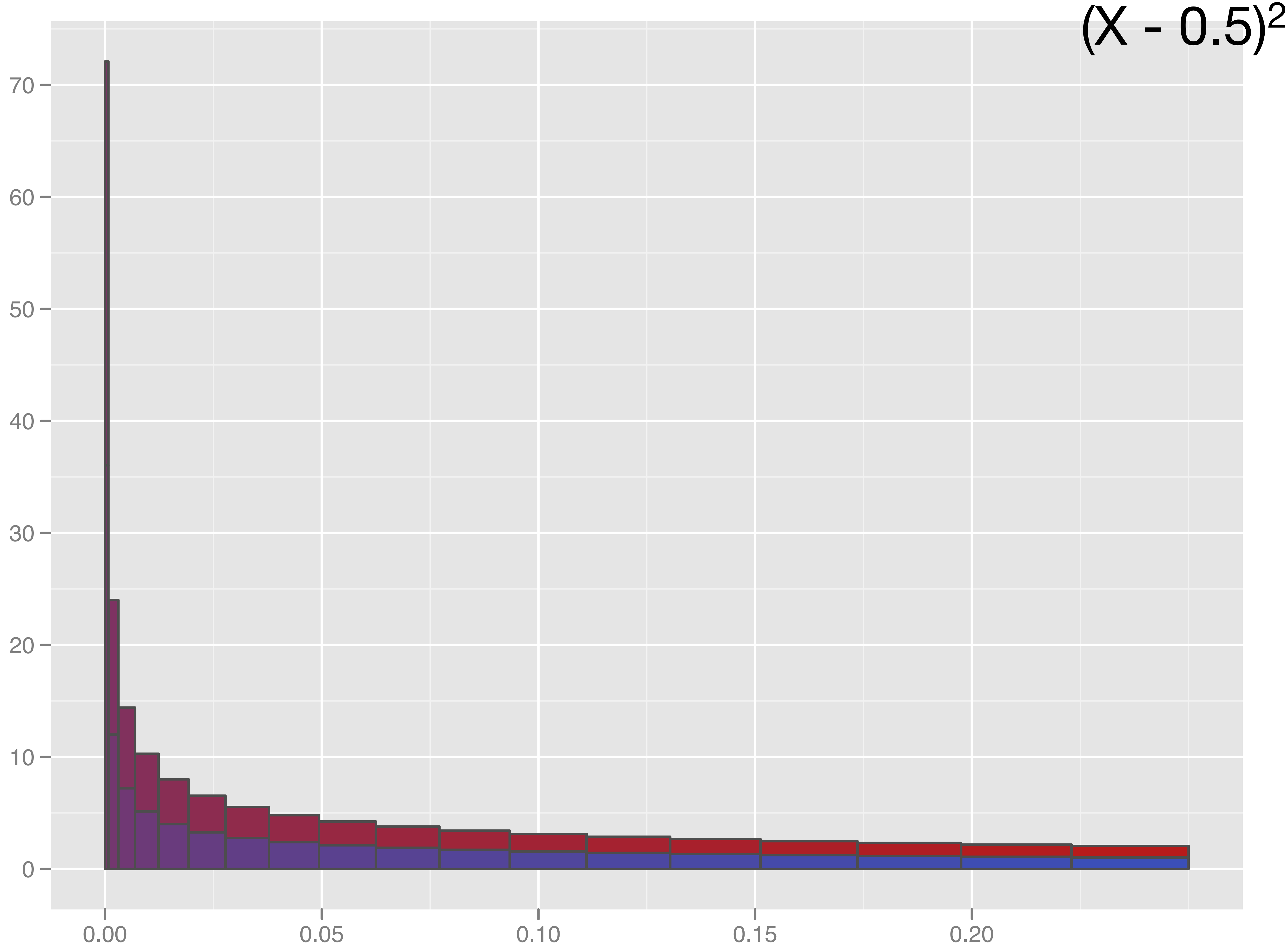
# |X - 0.5|



# Your turn

$X = \text{Uniform}(0, 1)$

$Y = (X - 0.5)^2$



$(X - 0.5)^2$
70
60
50
40
30
20
10
0
0.00
0.05
0.10
0.15
0.20

# Example

$A \sim \text{Exp}(\theta = 1)$

$B = 1 - e^{-A}$

Find $f_Y(y)$. Does y have a named distribution?

# Change of variables

# Change of variables

Transformation must have an inverse!

If $Y = u(X)$, and

$v$ is the inverse of $u$, $X = v(Y)$

then

$f_Y(y) = f_X(v(y))\ |v'(y)|$

# Your turn

$A \sim \text{Exp}(1)$

$B = 1 - e^{-A}$

Find $f_Y(y)$. Does y have a named distribution?

# Your turn

$X$ = Uniform(-1, 1)

$Y = X^2$

Can you use the change of variables technique here? Why/why not? If not, how could you modify X to make it possible?

# Relationship to uniform

Important connection between the uniform and every other random variable through the cdf.

# Uniform to any rv

IF

$Y \sim \text{Uniform}(0, 1)$

F a cdf

THEN

$X = F^{-1}(Y)$ is a rv with cdf $F(x)$

(Assume F strictly increasing for simplicity)

# Any rv to uniform

IF

X has cdf F

Y = F(X)

THEN

Y ~ Uniform(0, 1)

(Assume F strictly increasing for simplicity)