CS 237: Probability in Computing

Wayne Snyder
Computer Science Department
Boston University

Lecture 16:

- Normal Distribution continued;
- Standard Normal Distribution and Z-Scores;
- Properties of the Normal Distribution;
- The Normal Distribution as an approximation of the Binomial;
- The Yates Continuity Correction for continuous approximations of discrete distributions.

Normal Distribution

By using parameters to fit the requirements of probability theory (e.g., that the area under the curve is 1.0), we have the formula for the Normal Distribution, which can be used to approximate the Binomial Distribution and which models a wide variety of random phenomena:

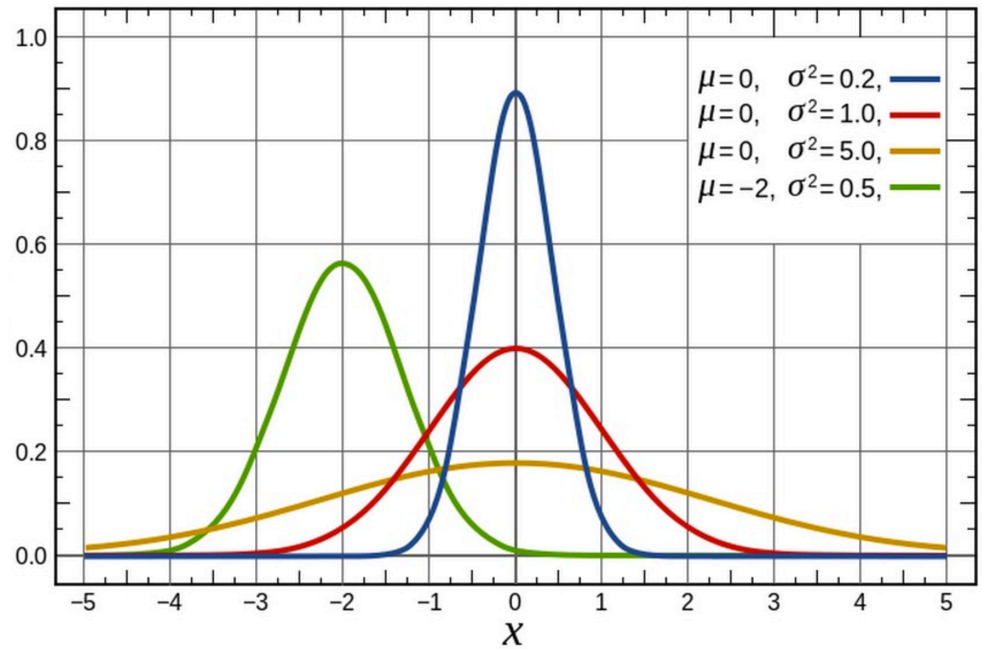
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where

μ = mean/expected value

σ = standard deviation

σ^2 = variance



Normal Distribution

Recall that the only way we can analyze probabilities in the continuous case is with the CDF:

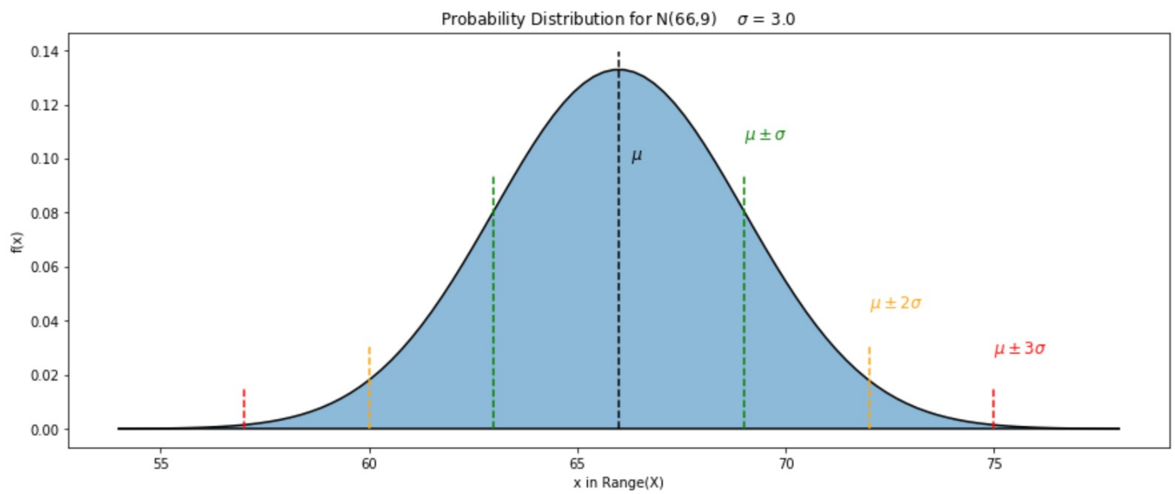
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

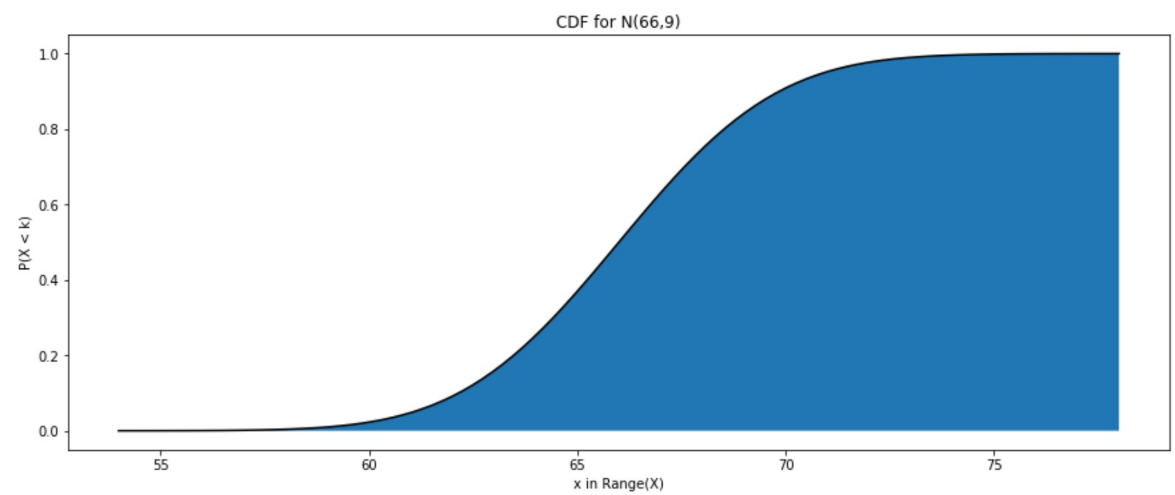
$$F(a) = \int_{-\infty}^{a} f(x)\, dx$$

P(X < a) = F(a)

P(X > a) = 1.0 – F(a)

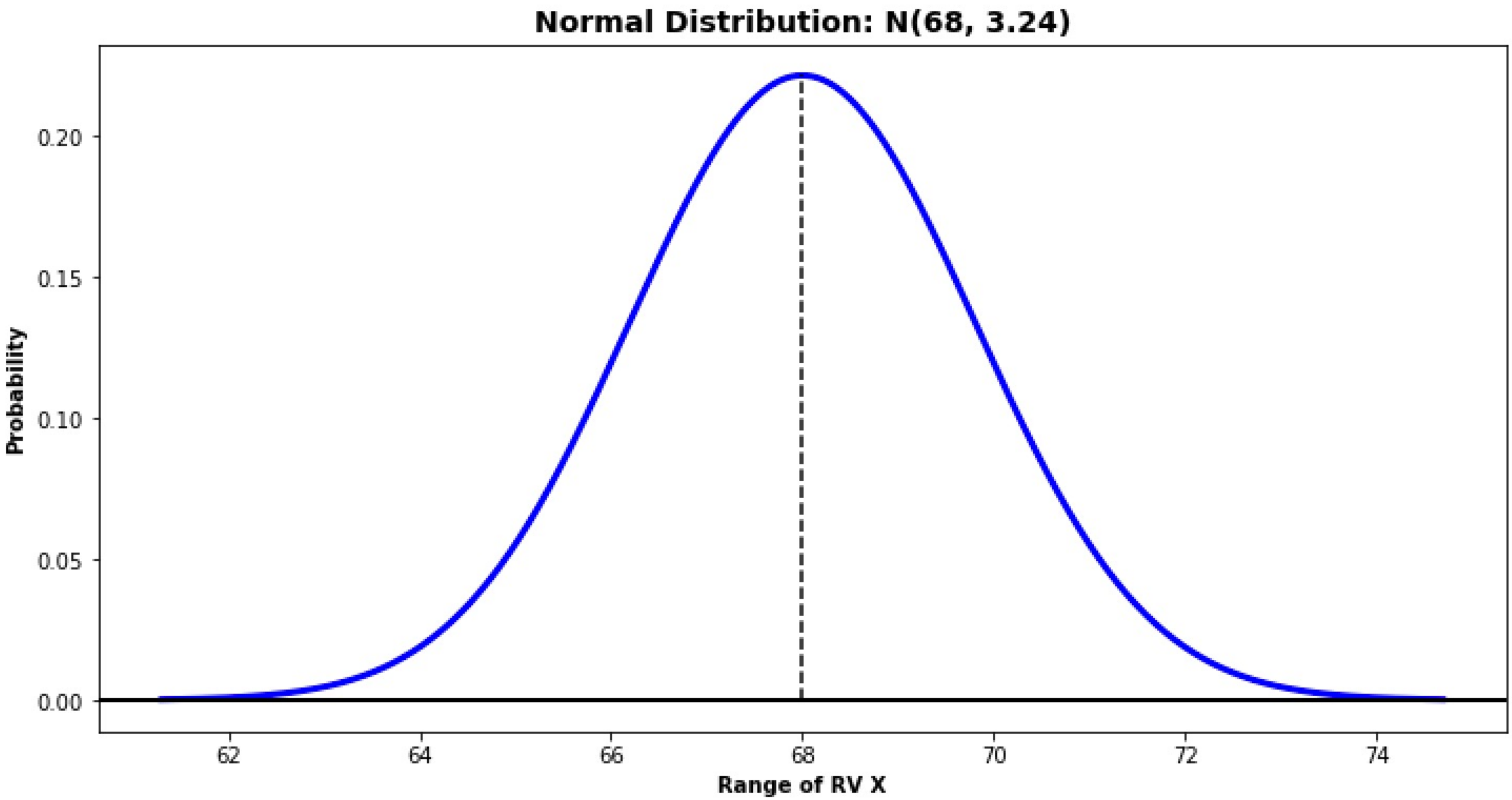
P(a < X < b) = F(b) – F(a)





Normal Distribution

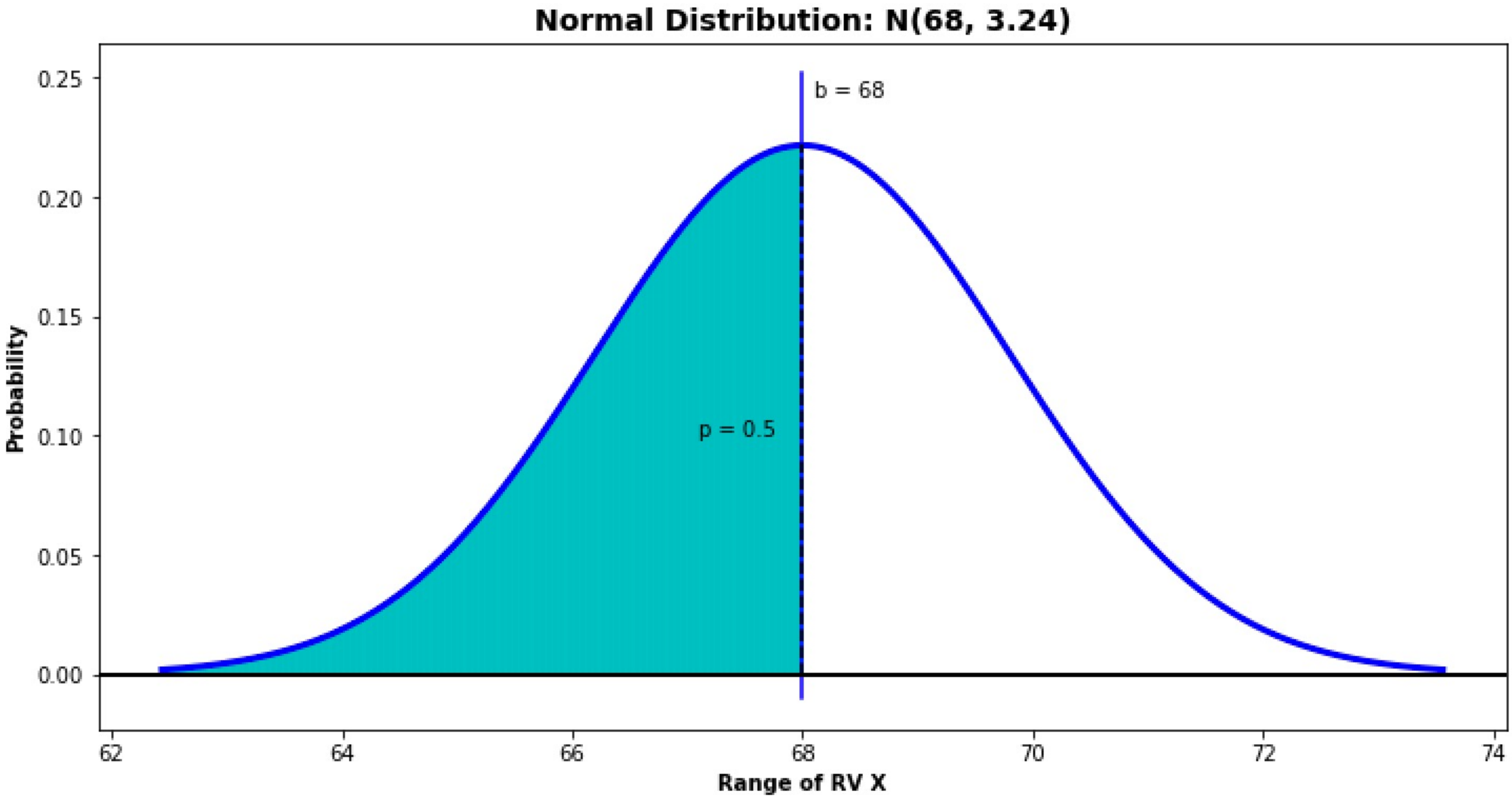
Suppose heights at BU are distributed normally with a mean of 68 inches and a standard deviation of 1.8 inches.




```
mean = 68
var = 3.24
stdev = 1.8
```

Normal Distribution

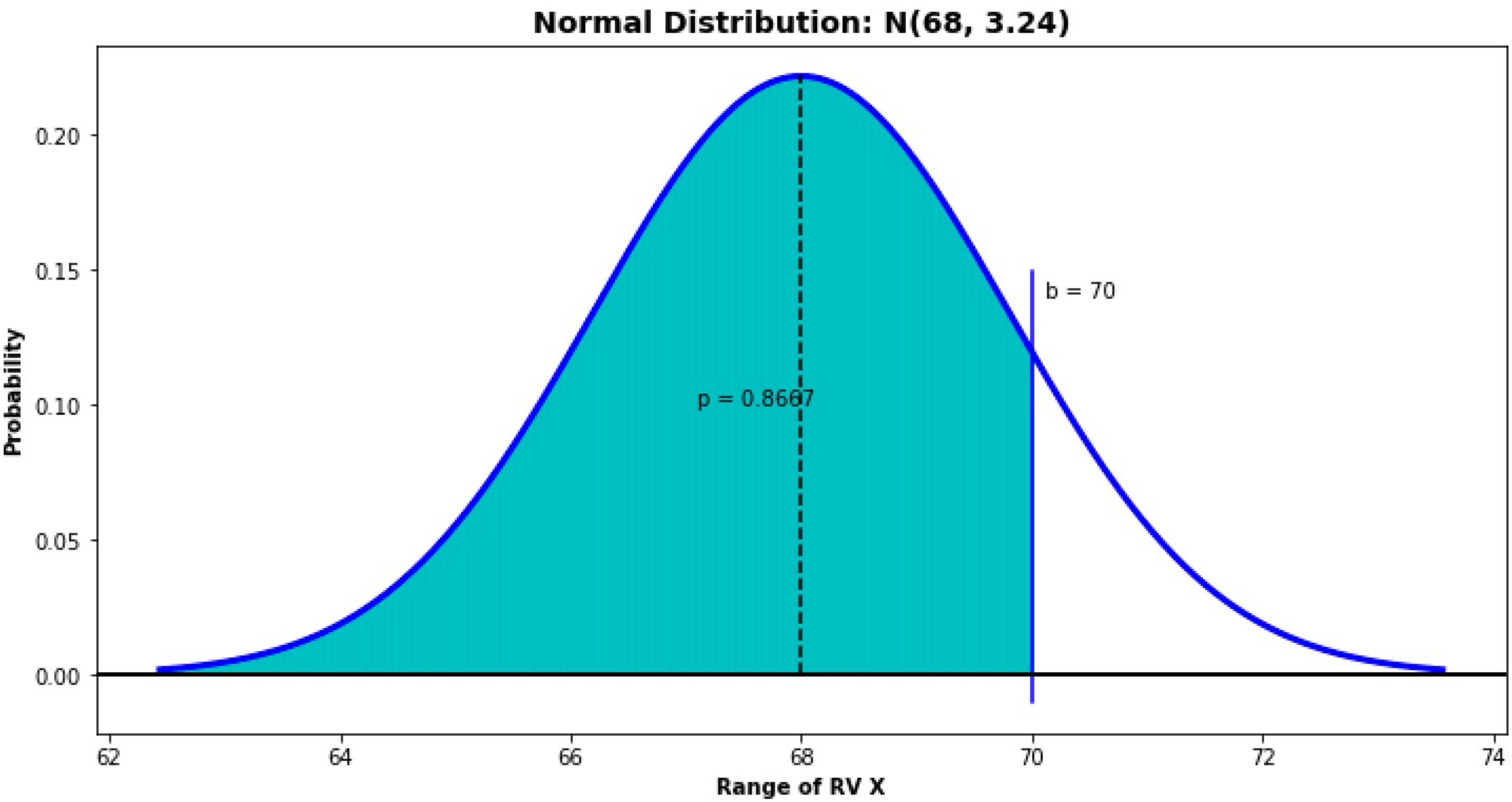
How many people are of less than average height?




```
mean = 68
var = 3.24
stdev = 1.8
P(X < 68) = 0.5
```

Normal Distribution

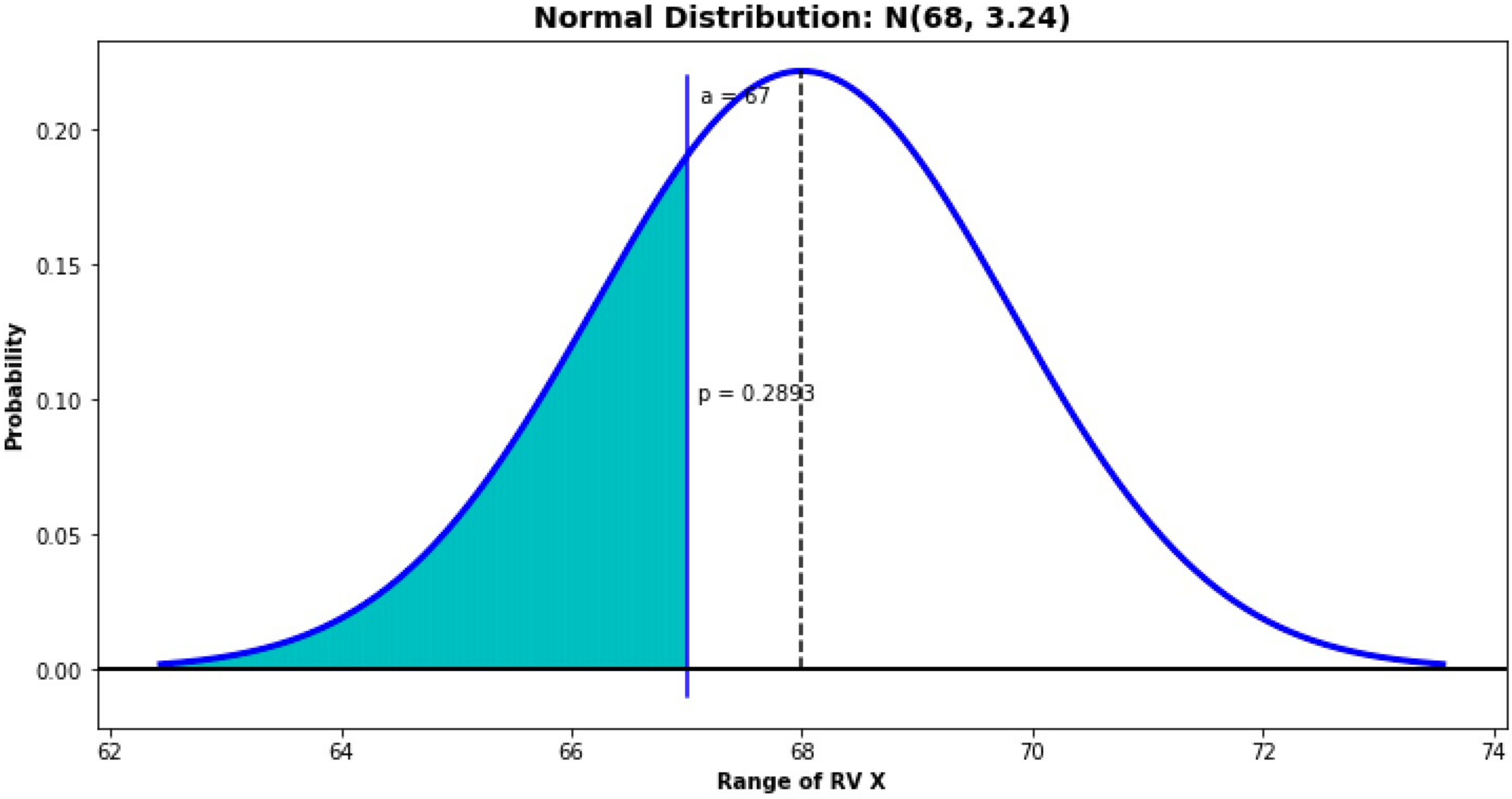
How many people are less than 70 inches?




```
mean = 68
var = 3.24
stdev = 1.8
P(X < 70) = 0.8667
```

Normal Distribution

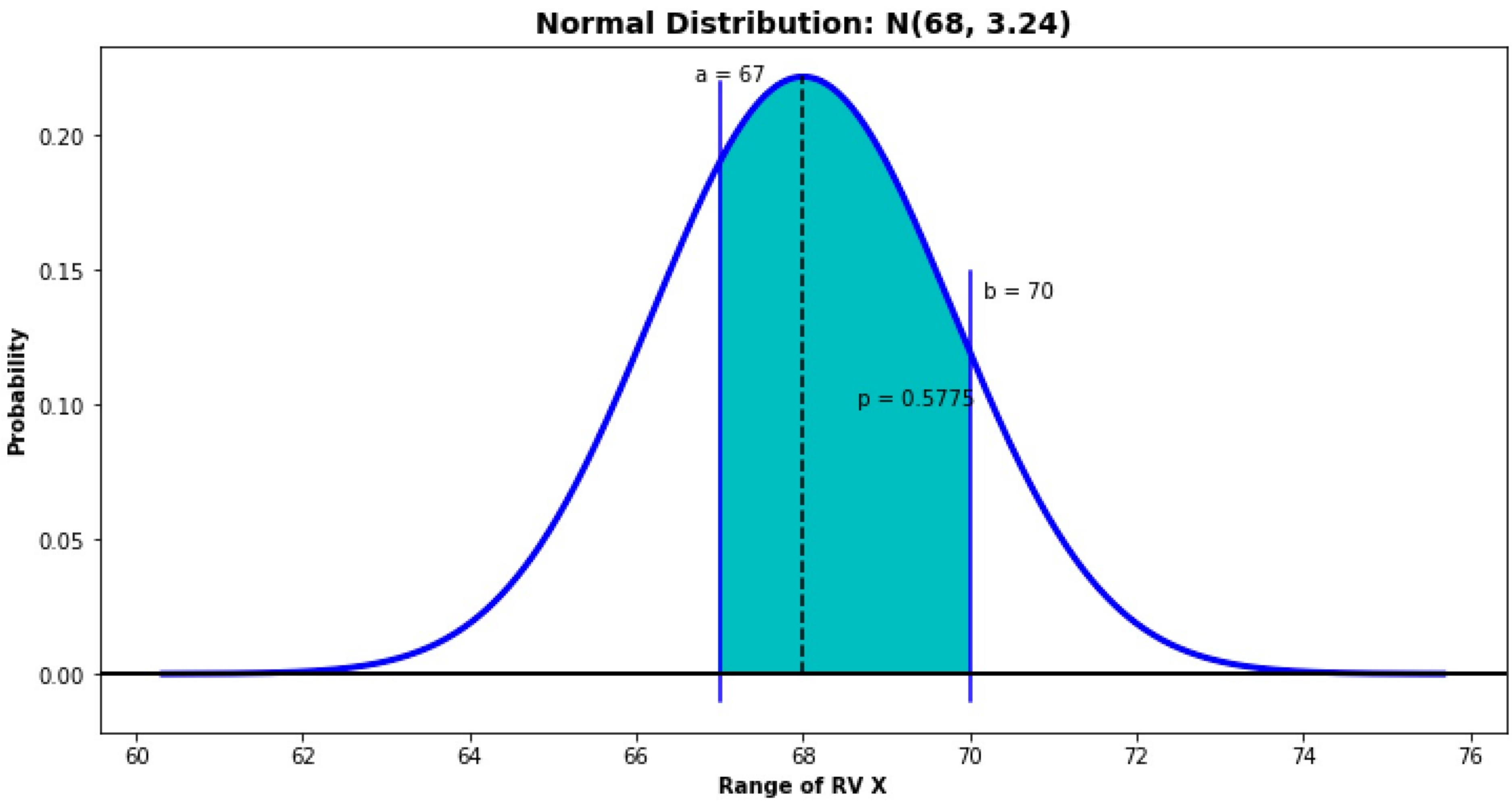
How many people are less than 67 inches?




```
mean = 68
var = 3.24
stdev = 1.8
P(X < 67) = 0.2893
```

Normal Distribution

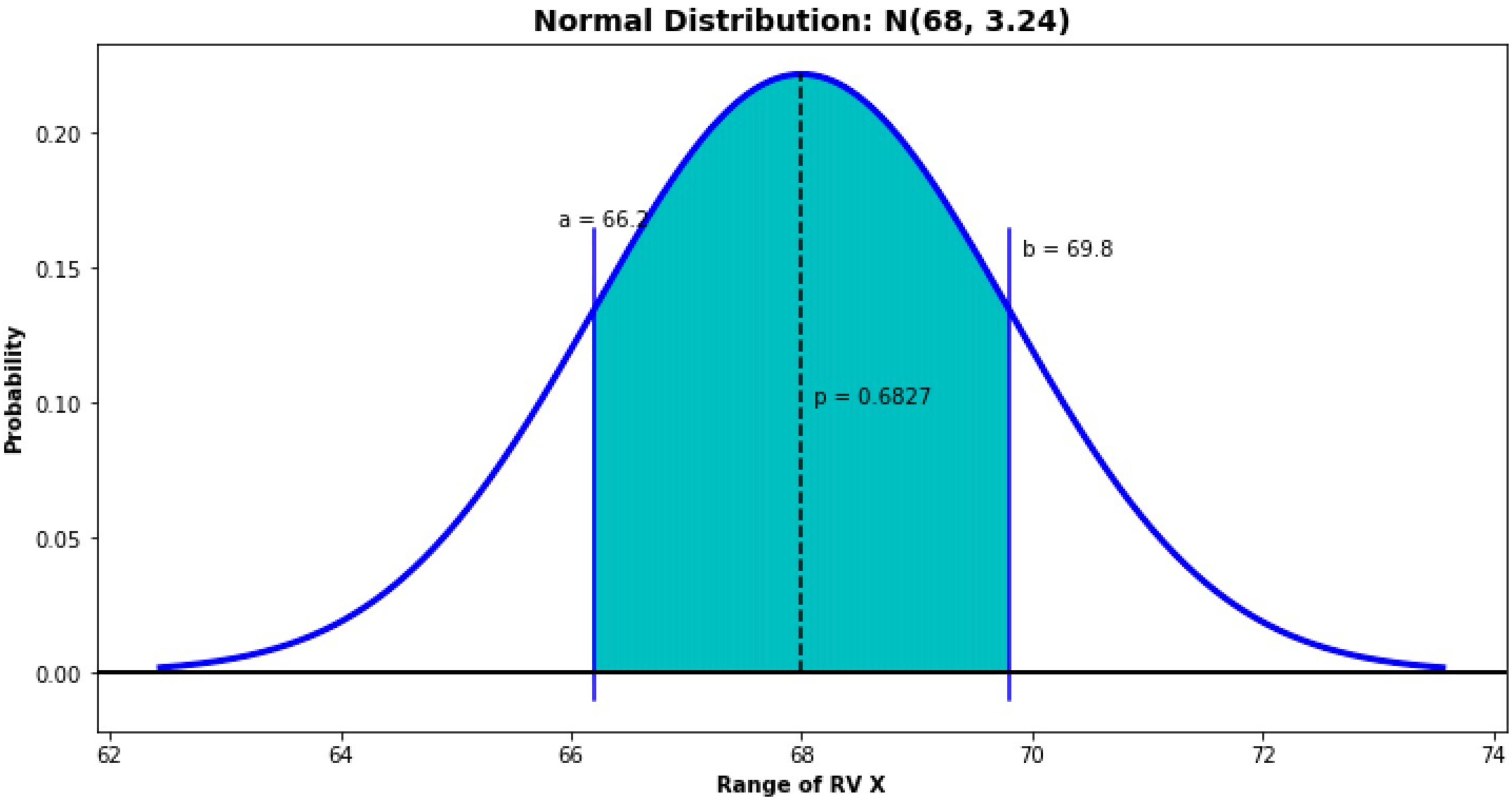
How many people are between 67 and 70 inches?




```
mean = 68
var = 3.24
stdev = 1.8
P(67 < X < 70) = P(X < 70) - P(X < 67) = 0.8667 - 0.2893 = 0.5775
```

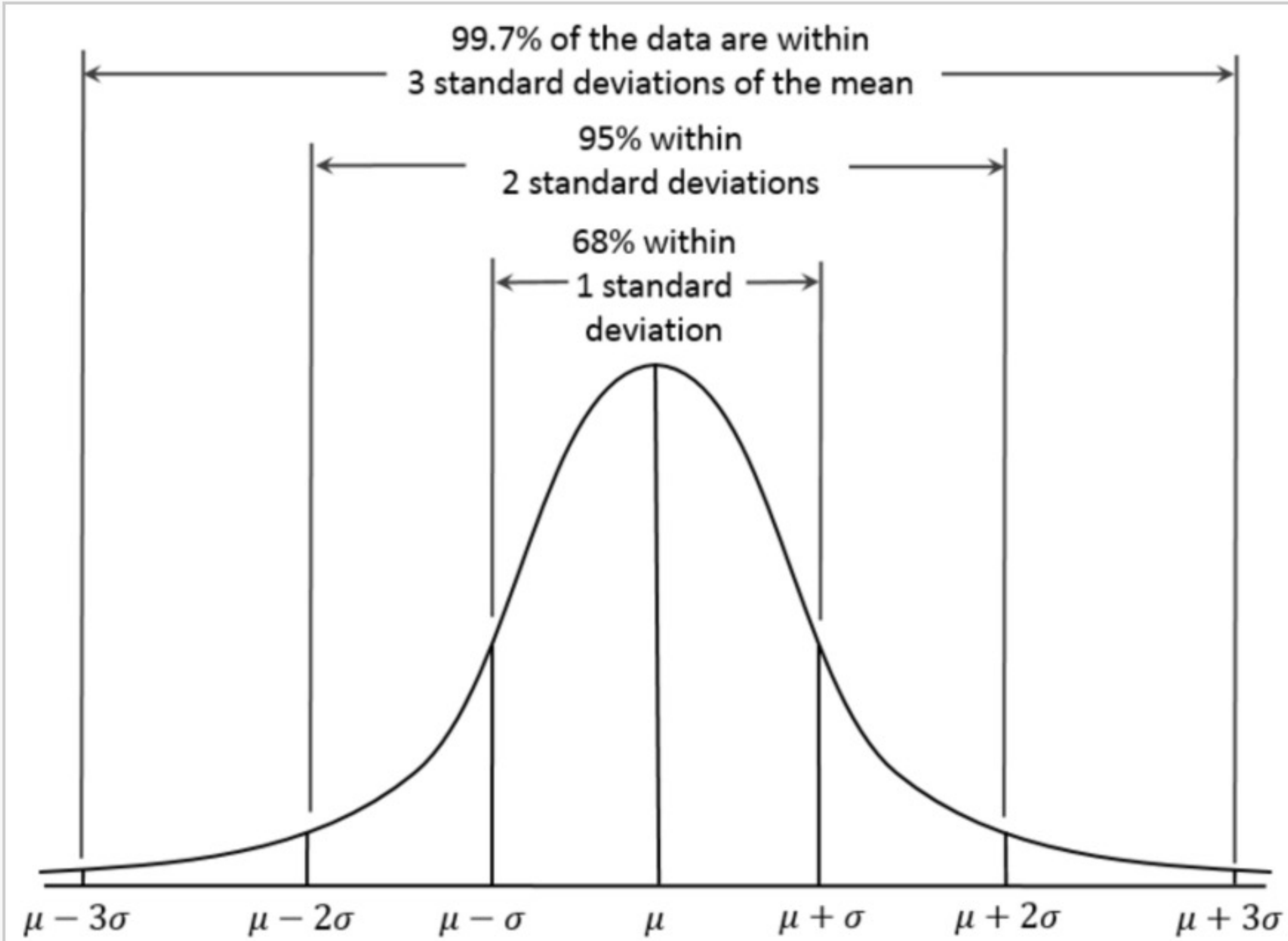
Normal Distribution

How many people are within one standard deviation of the mean height?




```
mean = 68
var = 3.24
stdev = 1.8
P(66.2 < X < 69.8) = P(X < 69.8) - P(X < 66.2) = 0.8413 - 0.1587 = 0.6827
```

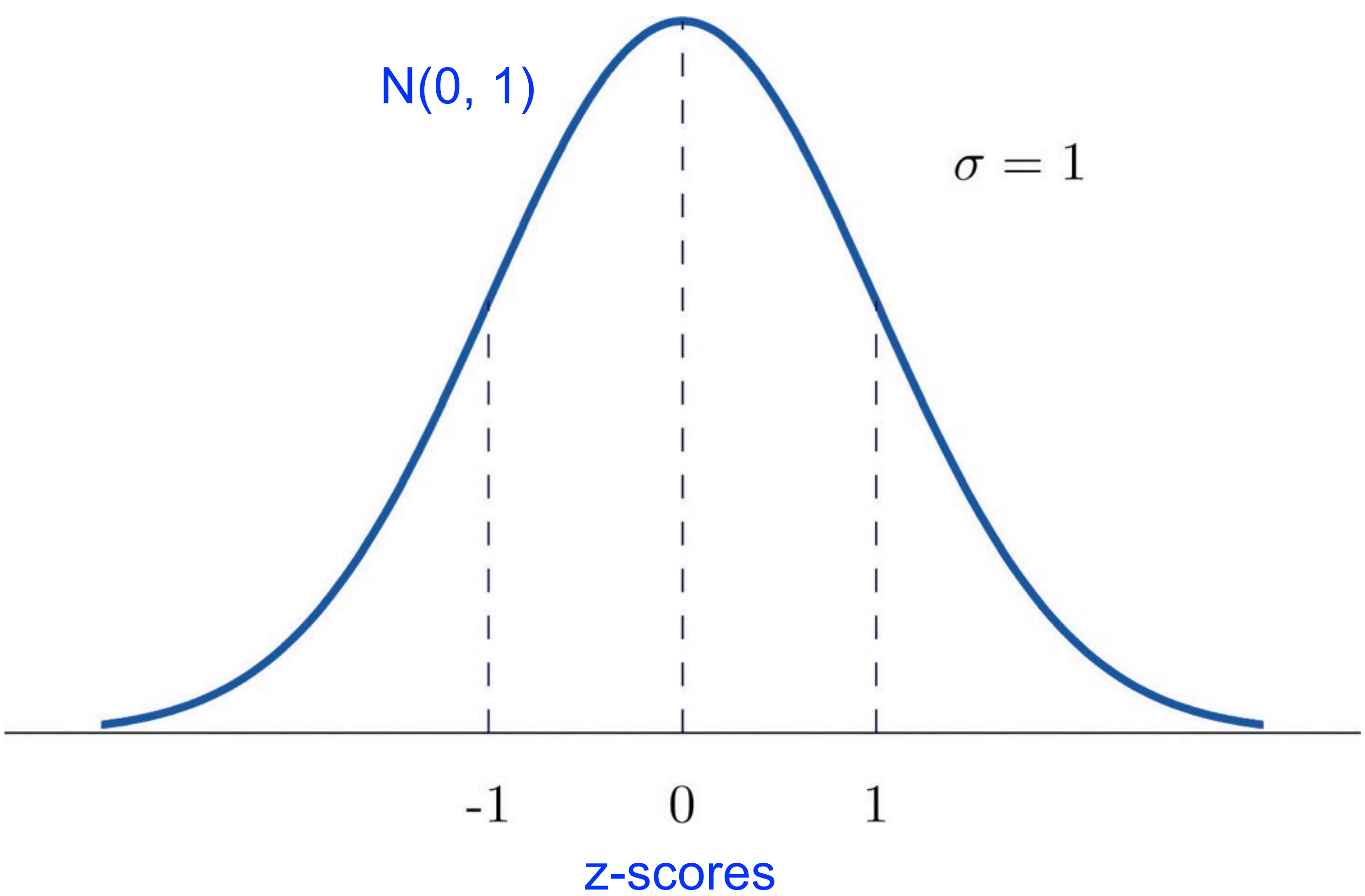
Normal Distribution



For the normal distribution, the values less than one standard deviation away from the mean account for 68.27% of the set; while two standard deviations from the mean account for 95.45%; and three standard deviations account for 99.73%.

Standard Normal Distribution

Since there are a potentially infinite number of Normal Distributions, usually we calculate using a normalized version, the Standard Normal Distribution with mean 0 and standard deviation (and variance) 1:

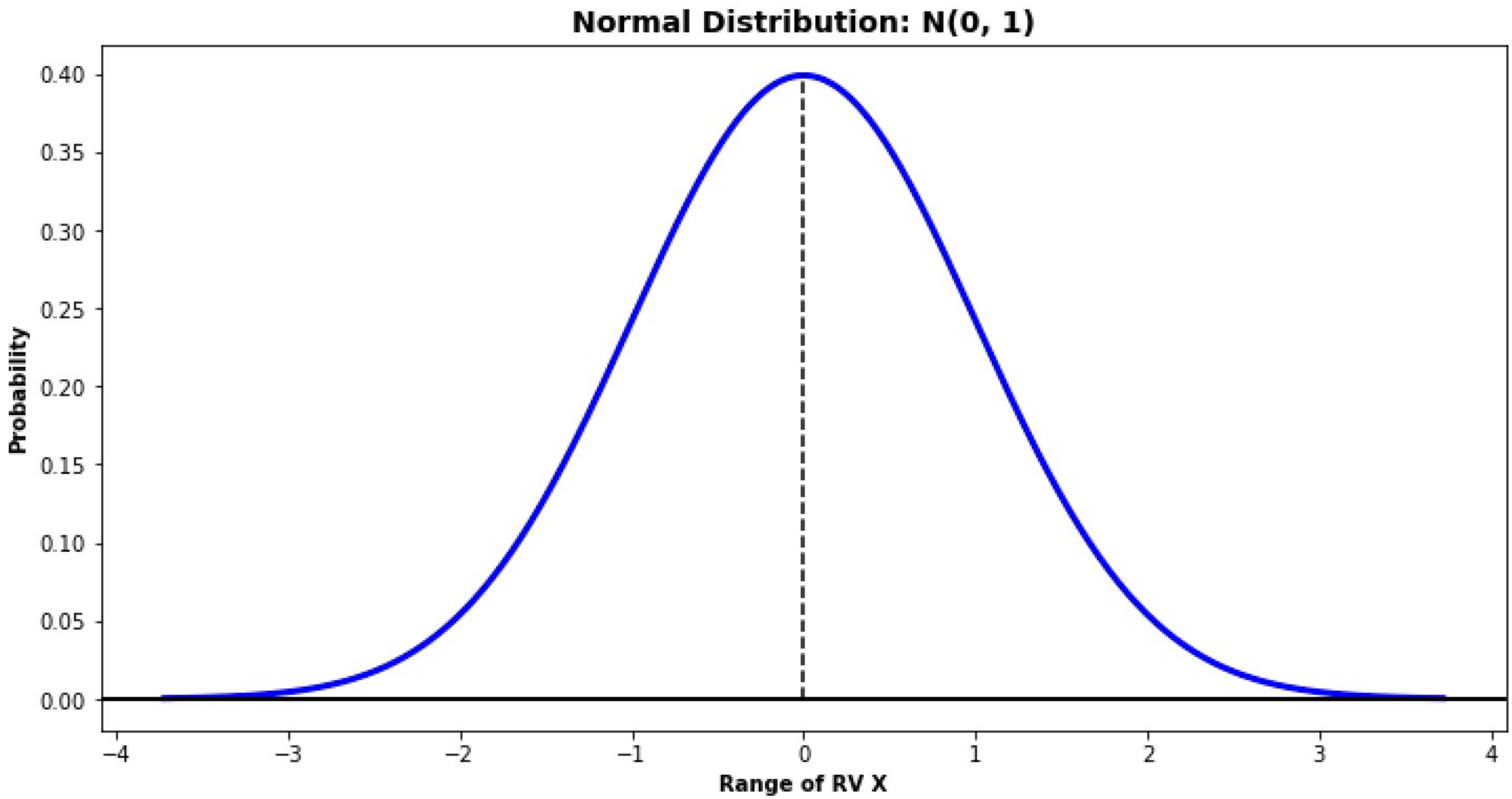


Any random variable X which has a normal distribution $N(\mu,\sigma^2)$ can be converted into a standardized random variable X* with distribution N(0,1) by the usual form. In statistics this always called Z:

$$Z = \frac{X - \mu}{\sigma}$$

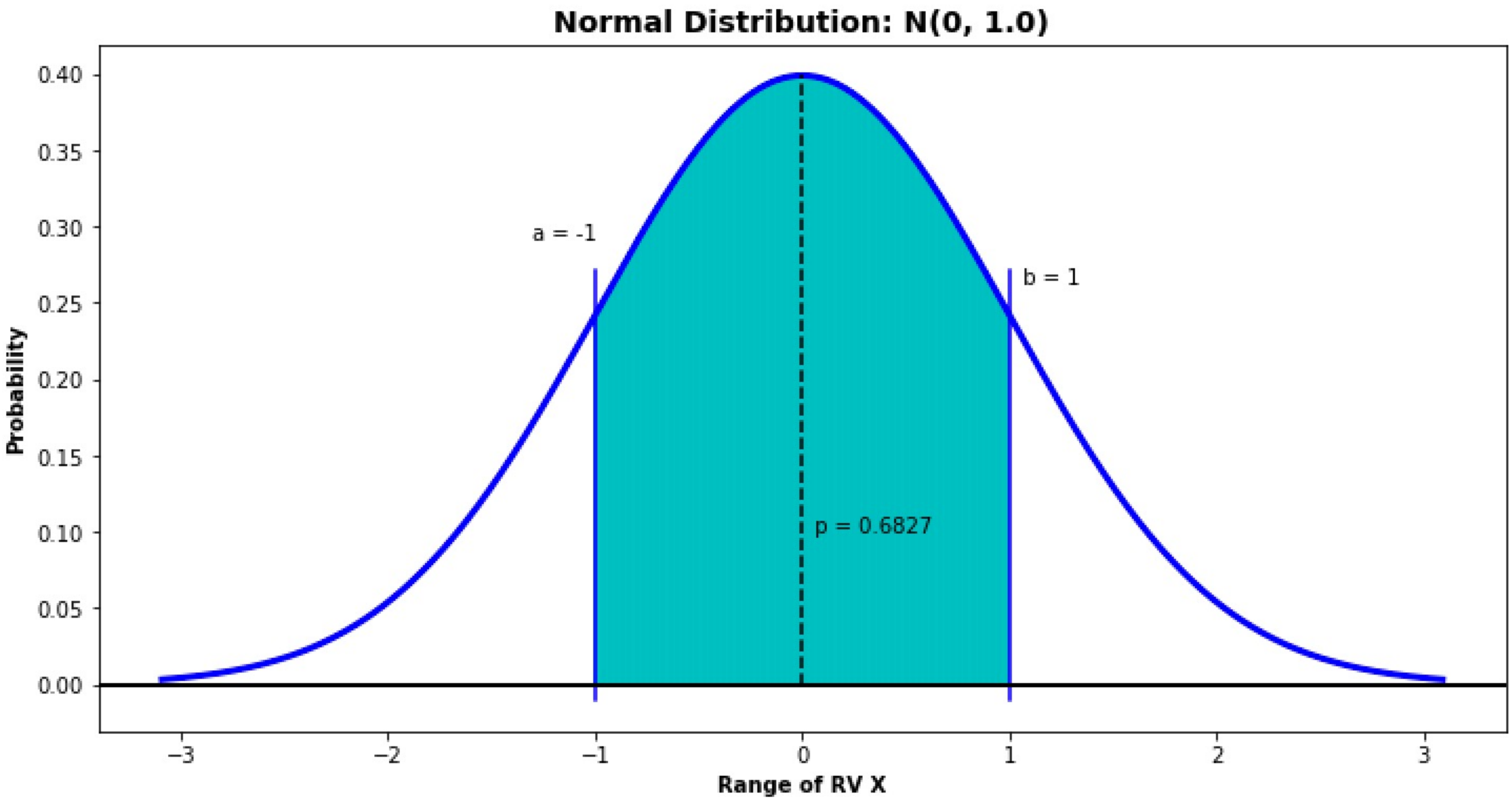
and the x-axis (the values returned by Z) are called z-scores.

Normal Distribution



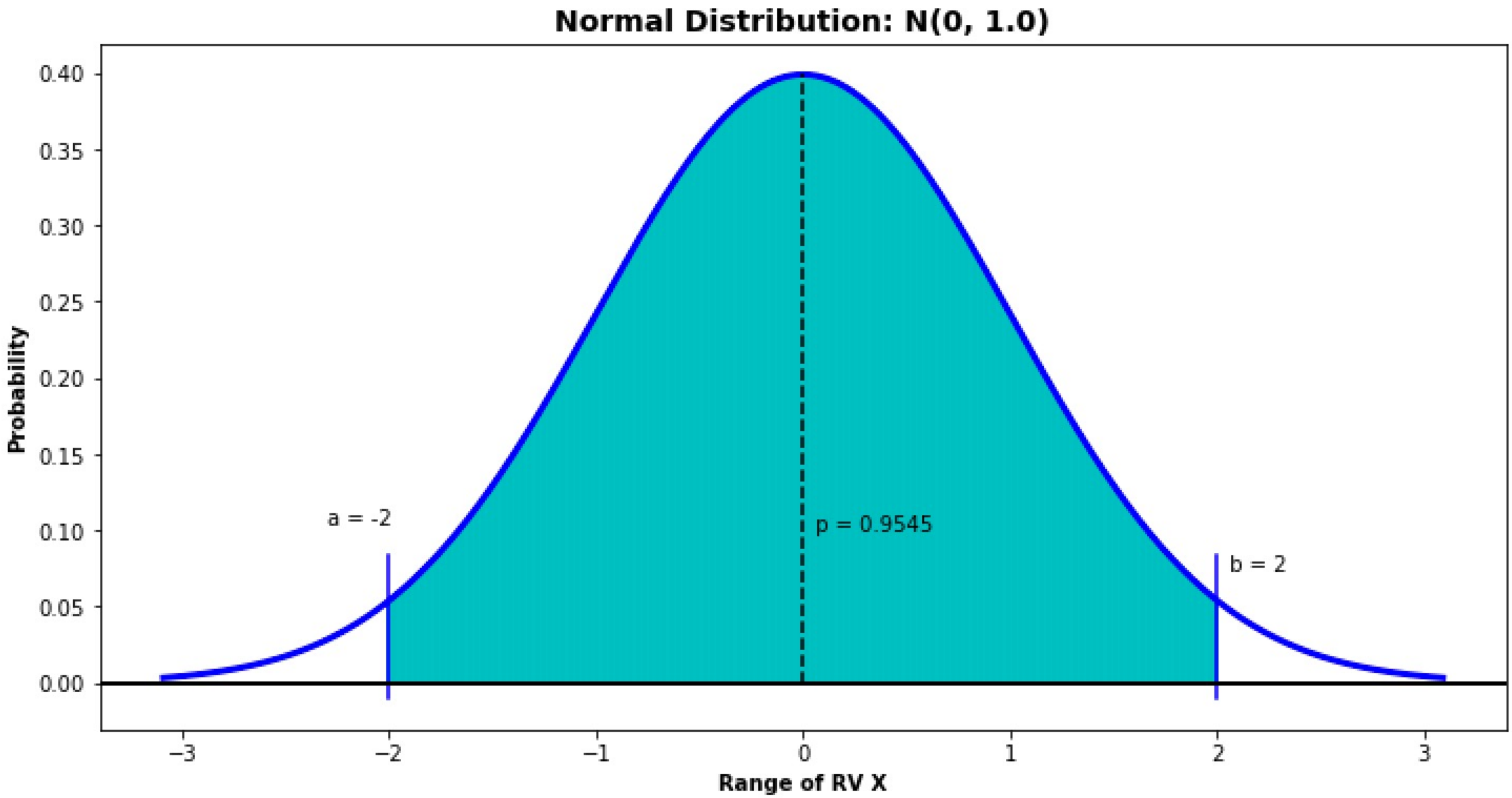

```
mean = 0
var = 1
stdev = 1
```

Normal Distribution



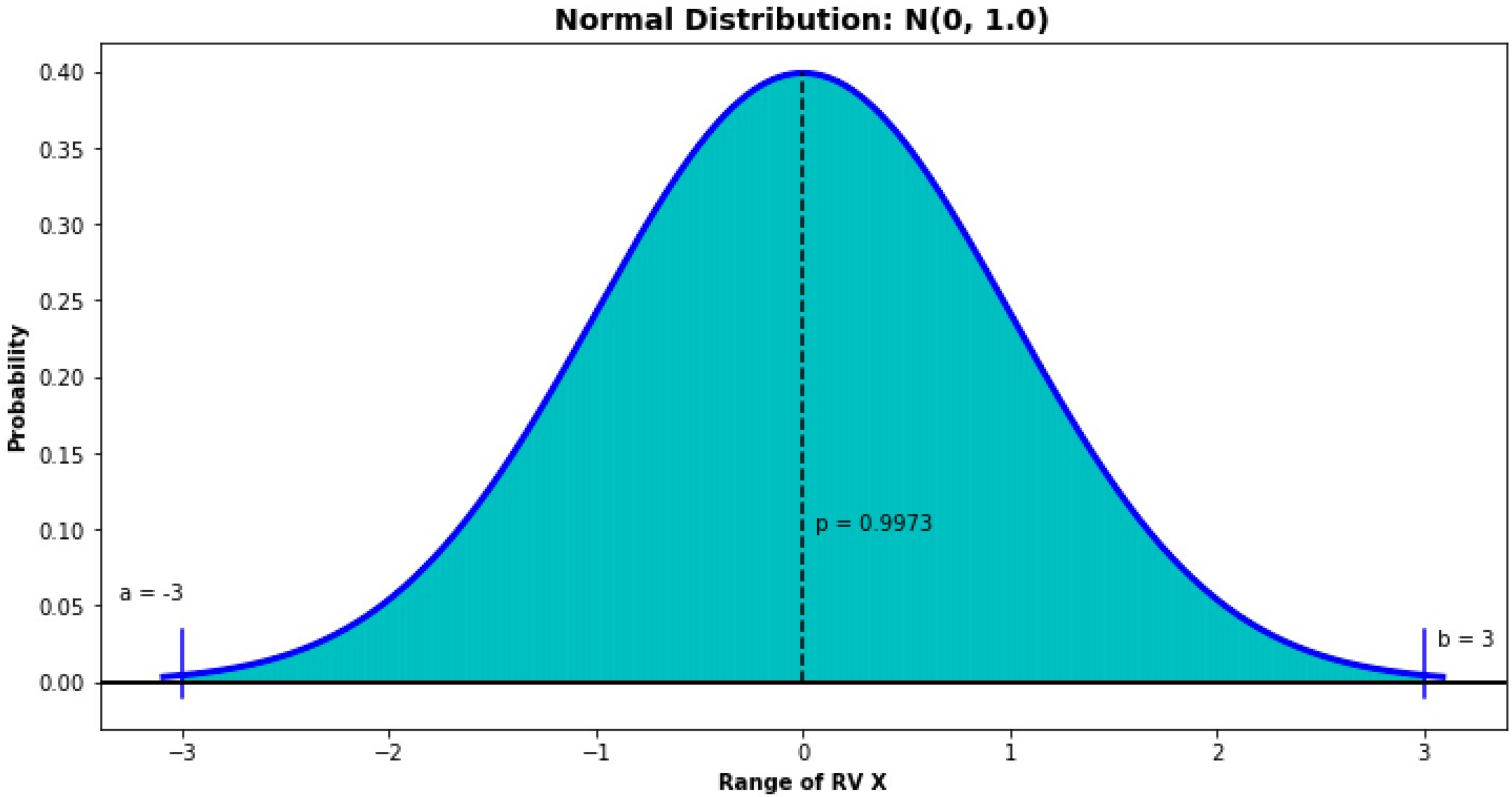

```
mean = 0
var = 1
stdev = 1.0
P(−1 < X < 1) = P(X < 1) − P(X < −1) = 0.8413 − 0.1587 = 0.6827
```

Normal Distribution




```
mean = 0
var = 1
stdev = 1.0
P(−2 < X < 2) = P(X < 2) − P(X < −2) = 0.9772 − 0.0228 = 0.9545
```

Normal Distribution




```
mean = 0
var = 1
stdev = 1.0
P(−3 < X < 3) = P(X < 3) − P(X < −3) = 0.9987 − 0.0013 = 0.9973
```

Properties of the Normal Distribution

The normal distribution has interesting properties, which follow directly from our previous results on mean and variance:

The sum of independent normally-distributed random variables is still normal:

If $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ with X and Y independent, then

$$X + Y \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2).$$

The linear transformation of a normally random variables is still normal:

If $X \sim N(\mu, \sigma^2)$ and $Y = aX + b$ then

$$Y \sim N(a\mu + b, a^2\sigma^2)$$

where $\sigma_Y = |a|\sigma$.

The normal, like the binomial, is useful in any situation where the outcome depends on the sum of a large number of Bernoulli RVs (binary choices):

- Bought this stock or not;
- This gene is dominant or recessive;
- Got delayed at a traffic light or not;

In fact, it is quite usual to use the normal as an approximation for the binomial.....

Normal Approximation to the Binomial

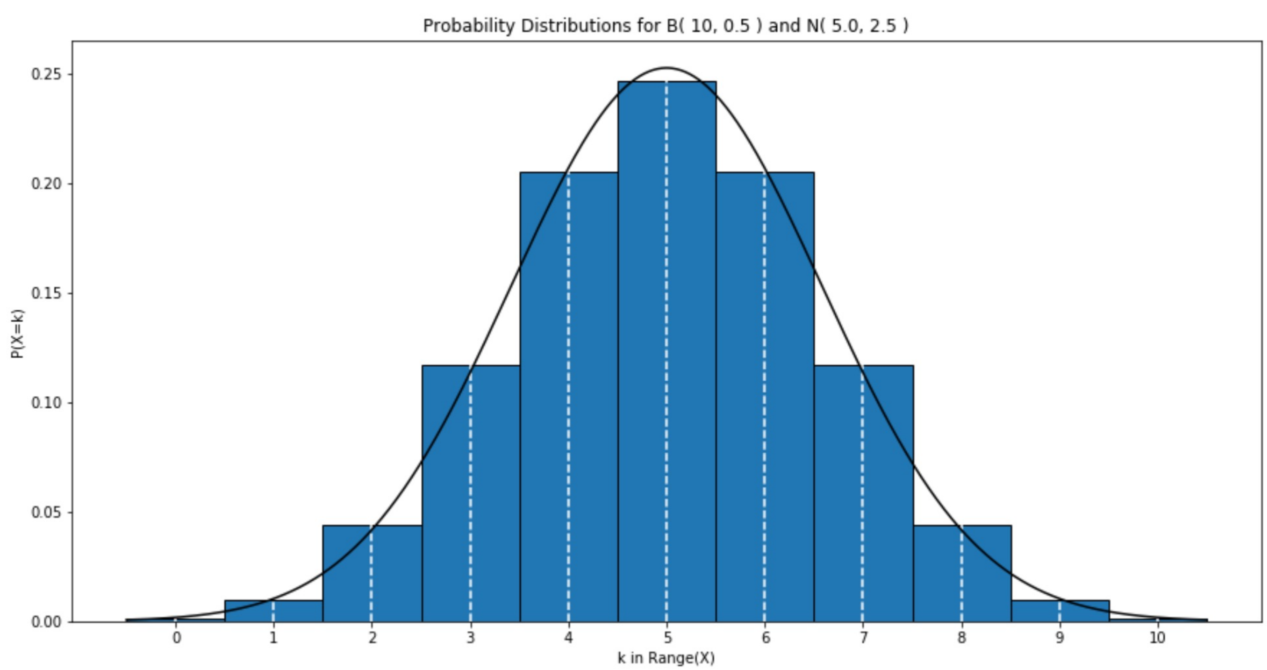
Since the binomial distribution with p close to 0.5 can be viewed as the limit of the binomial as N gets large, we can use the normal distribution to approximate the binomial.

To do the approximation, we have to have the same mean and standard deviation, so for a binomial RV

$$X \sim B(n, p)$$

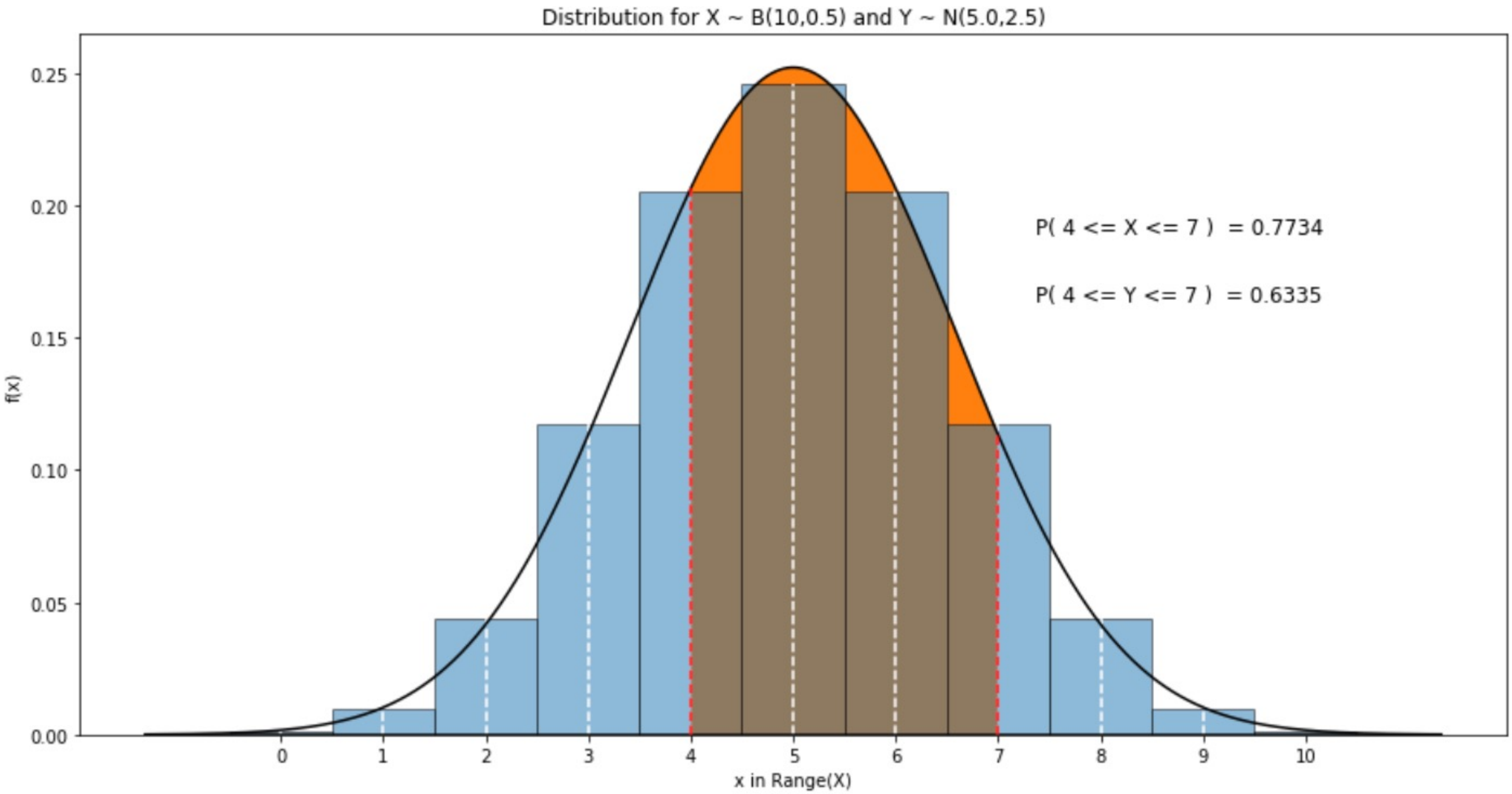
we consider a normal RV

$$Y \sim N(np, np(1-p))$$



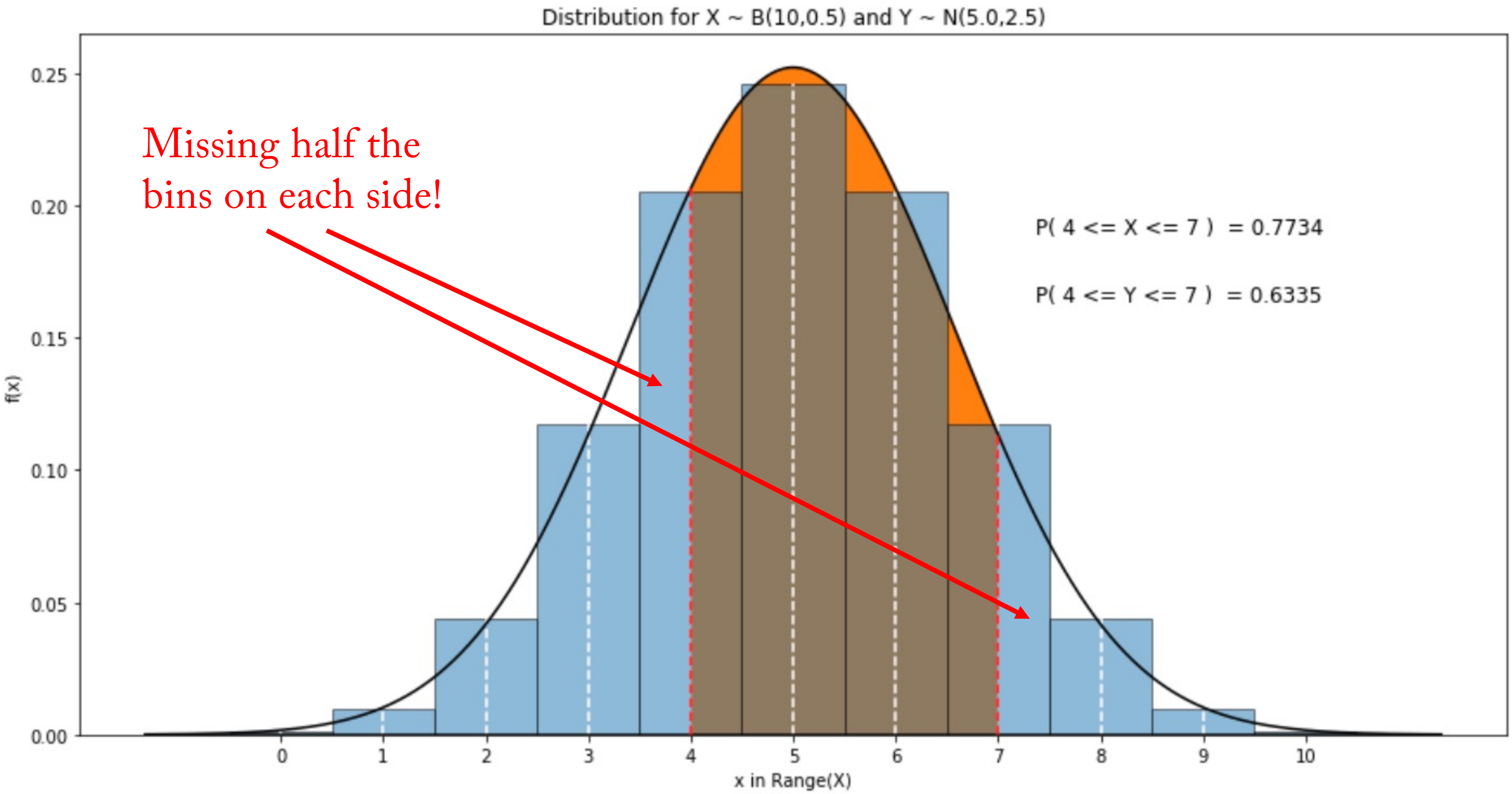
Normal Approximation to the Binomial

However, this is not quite right when we come to do actual calculations:



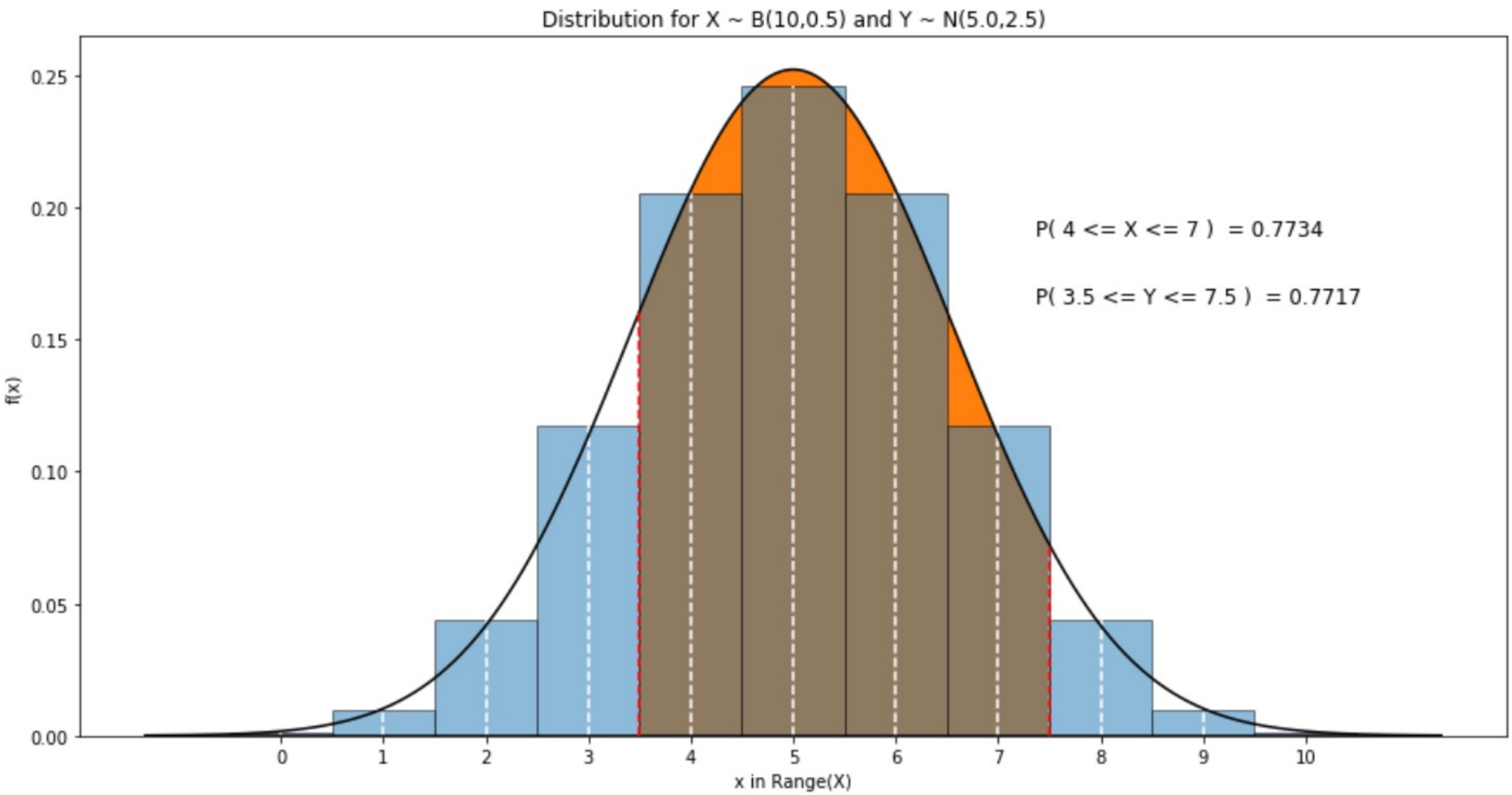
Normal Approximation to the Binomial

However, this is not quite right when we come to do actual calculations, as we underestimate by stopping half way through the lower and upper bins!



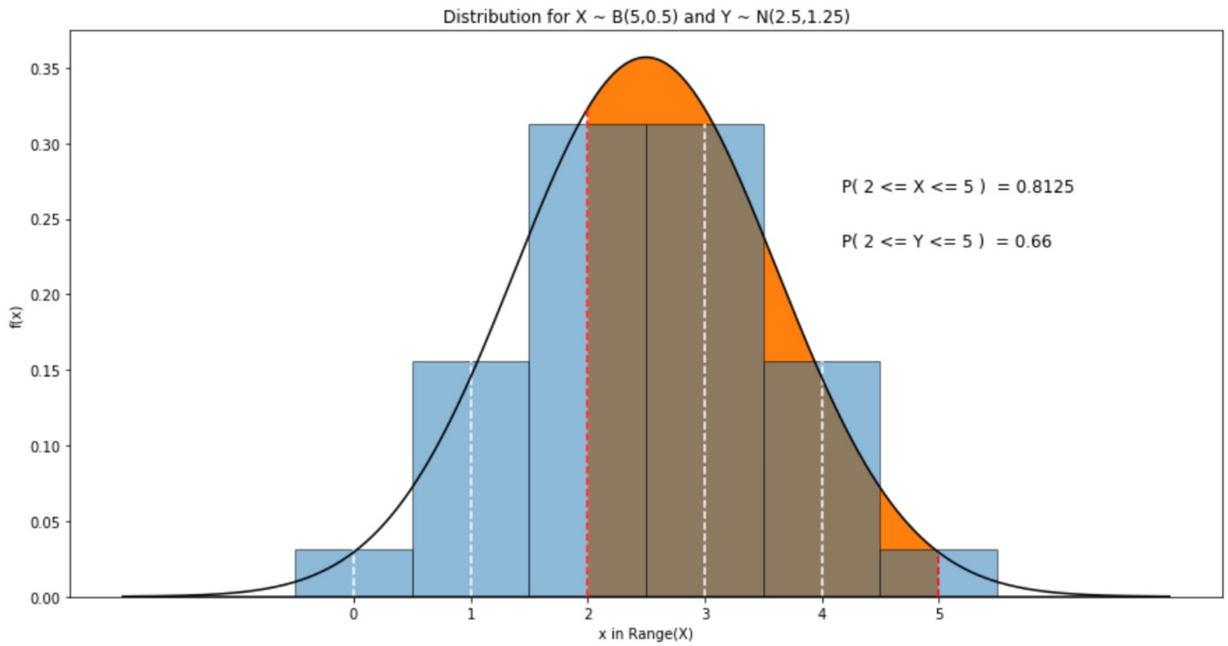
Normal Approximation to the Binomial

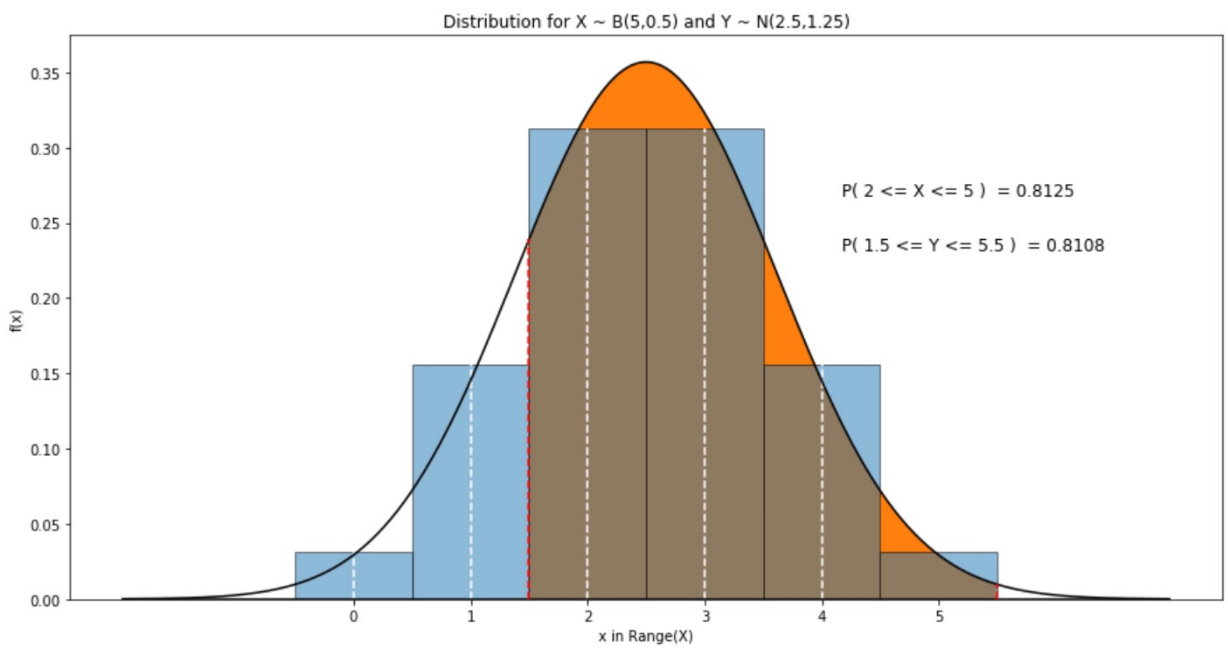
So, the continuity correction is to subtract 0.5 from any lower bound and add 0.5 to any upper bound:



Normal Approximation to the Binomial

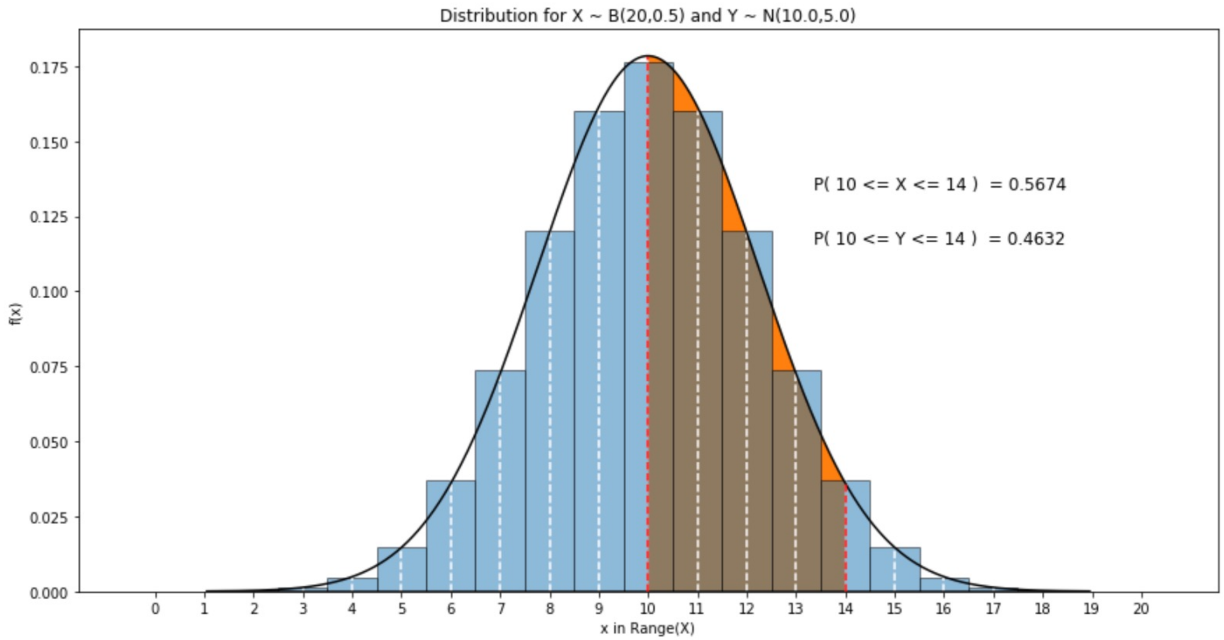
This actually works well even for small N:

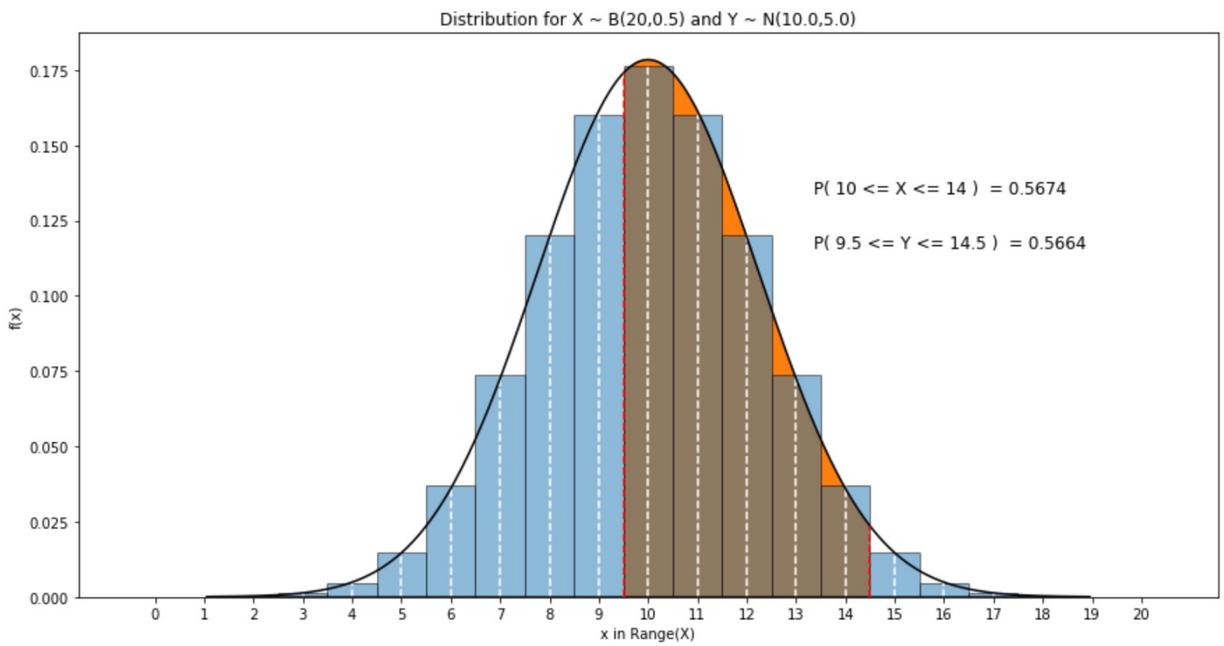




Normal Approximation to the Binomial

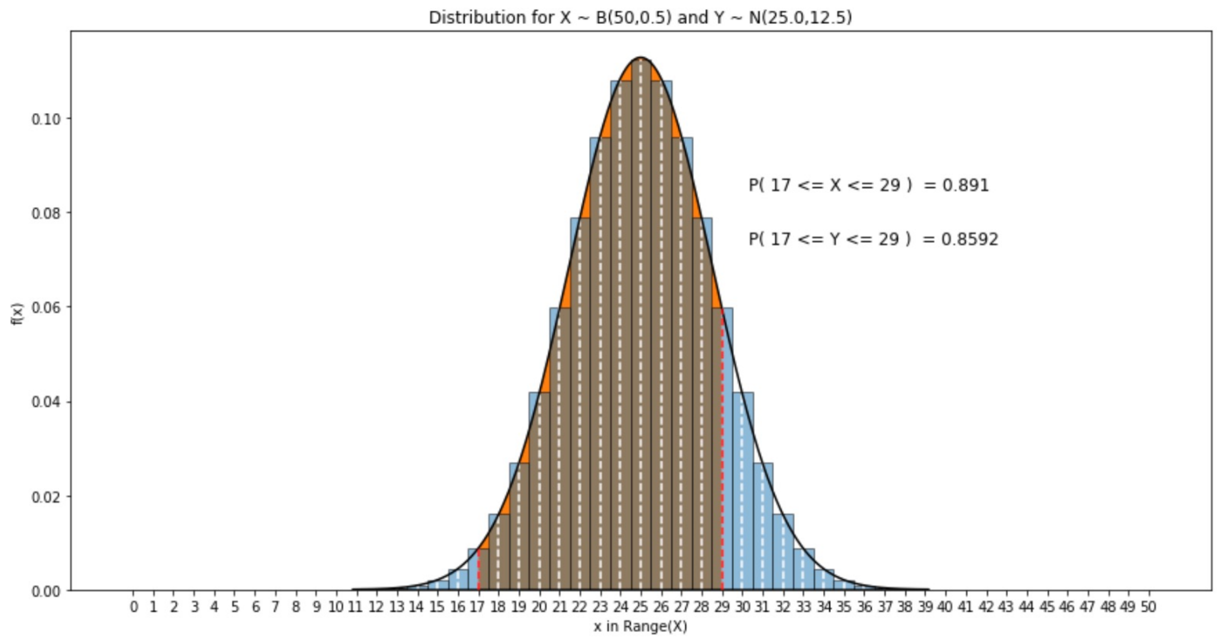
For larger N the uncorrected value gets better, but the correction is **still** better:

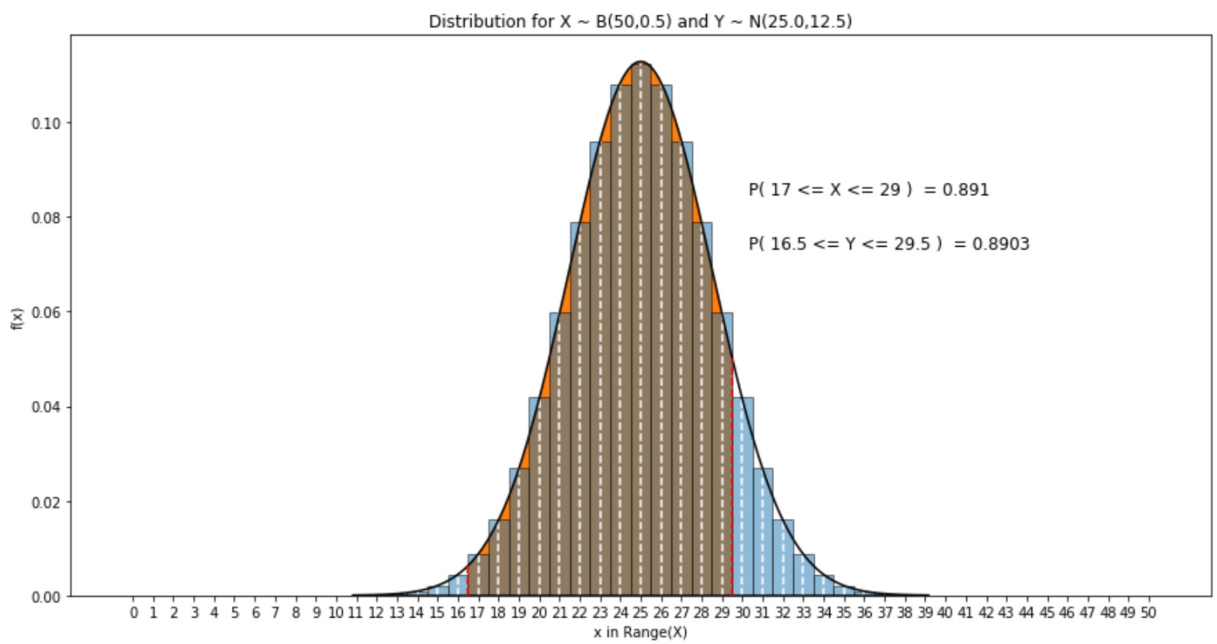




Normal Approximation to the Binomial

For larger N the uncorrect value gets better, but the correction is still better:





Normal Approximation to the Binomial

For larger N the uncorrect value gets better, but the correction is still better: