The only quantities governing the binomial distribution are n, the number of experiments and π, the probability of the event of interest occurring in a given experiment. Therefore, all the properties of a binomial random variable S are function of n and π. The expected value and standard deviation of $S \sim Bin(n, \pi)$ are

Var(S) = n×π×(1−π)

$$E(S) = n\pi \quad \text{and} \quad \text{SD}(S) = \sqrt{n\pi(1-\pi)} \tag{3.8}$$

For example, if we observe 100 experiments and in each one the probability of A is 0.25, we expect 25 occurrences of A. The form of the standard deviation may seem a little strange but, after a little reflection, it should make sense. The standard deviation is a measure of variation. Suppose π is very close to 0. Then A almost never occurs. Therefore, S is almost always 0; that is, there is very little variation in S. The same argument applies if π is very close to 1, except that A almost always occurs and S is almost always n. That is, when π is close to either 0 or 1, then the standard deviation should be small. We expect a lot of variation whenever $\pi = 1/2$ because A and "not A" are equally likely.

3.9 The Normal Distribution

continuous random variable
(Bell shaped distribution)

The second important distribution that we will consider is the *normal distribution*. Unlike the binomial distribution, the normal distribution is a continuous distribution, and if a random variable X has a normal distribution, X can take any value between $-\infty$ and ∞, although extreme values are unlikely.

The normal distribution is governed by two parameters, traditionally denoted by μ and σ. Here, μ represents the mean of the distribution of X, and σ represents the standard deviation; because standard deviations are always positive, $\sigma > 0$. We write

$$\boxed{X \sim N(\underbrace{\mu}_{\text{mean}}, \overbrace{\sigma}^{\text{Standard deviation}})}$$

The shape of the distribution is given by the well-known bell-shaped curve, which takes its maximum value at μ; σ governs how spread out the curve is. Figure 3.3 shows a few normal distribution, corresponding to different values of μ and σ. These plots illustrate some important properties of the normal distribution. For instance, the distribution is symmetric about its peak, which occurs at the mean of the distribution. When the value of μ changes, the effect on the distribution is a shift; other aspects of the distribution, such as it "bell-shape," don't change. When the value of σ changes, the effect is essentially to change the scale on the x-axis.

Although it is easy to describe the shape of the normal distribution, it is a lite more difficult to determine probabilities associated with a normal distribution. Let X denote a random variable with a normal distribution with mean μ and standard deviation σ. Since X is a continuous random variable, we can't give a table listing the possible values of X together with their probabilities. Instead, we consider the probability that X falls into a certain range of values.

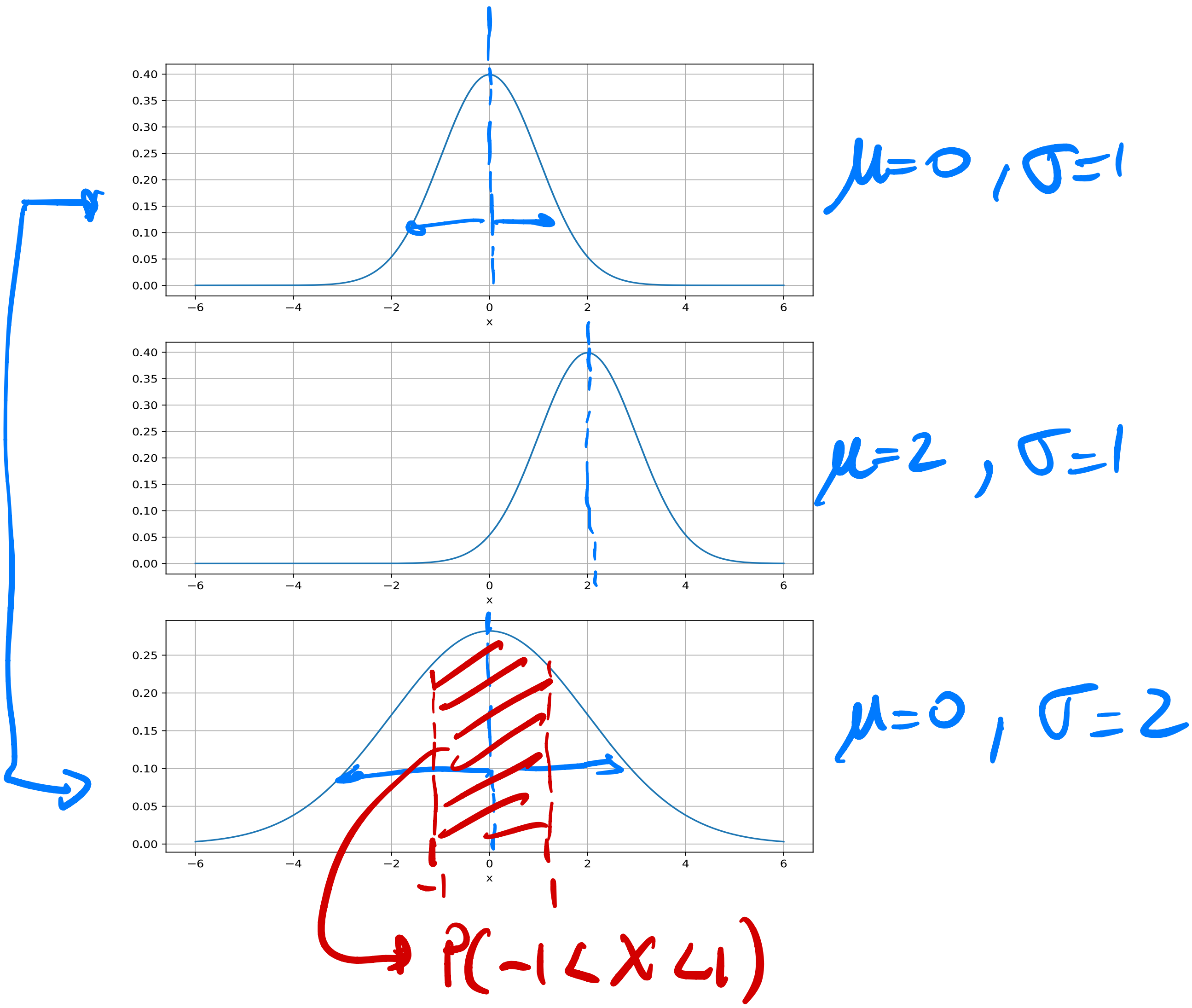


Figure 3.3: Example of normal distributions

To describe such probabilities, it is useful to relate X to a *standard normal distribution.* A standard normal distribution is the one in which $\mu = 0$ and $\sigma = 1$. Let X denote a random variable with mean μ and standard deviation σ. We can convert X to a random variable with a standard normal distribution by computing:

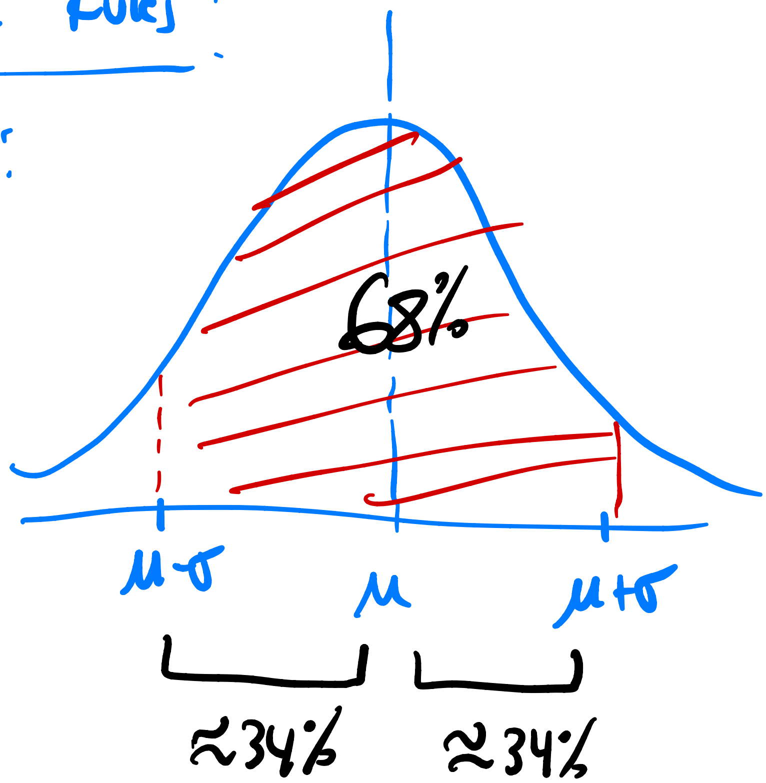
$$Z = \frac{X - \mu}{\sigma} \tag{3.9}$$

Then Z has a standard normal distribution. The standard normal distribution is a convenient reference distribution that can be used to understand variation in many different contexts. Let Z have a standard normal distribution and consider $P(-a < Z < a)$ as a function of a.

standard normal random variable ($\mu=0, \sigma=1$)
$\Rightarrow N(0, 1)$

Empirical Rules:

1st Rule:



Note that $P(-a < Z < a)$ rapidly approaches 1 as a increases. This is an important property of the normal distribution; with high probability, normal random variables tend to be close to their mean value.

Now, consider a normal random variable X that has mean μ and standard deviation σ. To find $P(-b < X < b)$ for some value b, we convert this probability into a probability concerning a standard normal random variable Z. Because $\frac{X-\mu}{\sigma}$ has a standard normal distribution

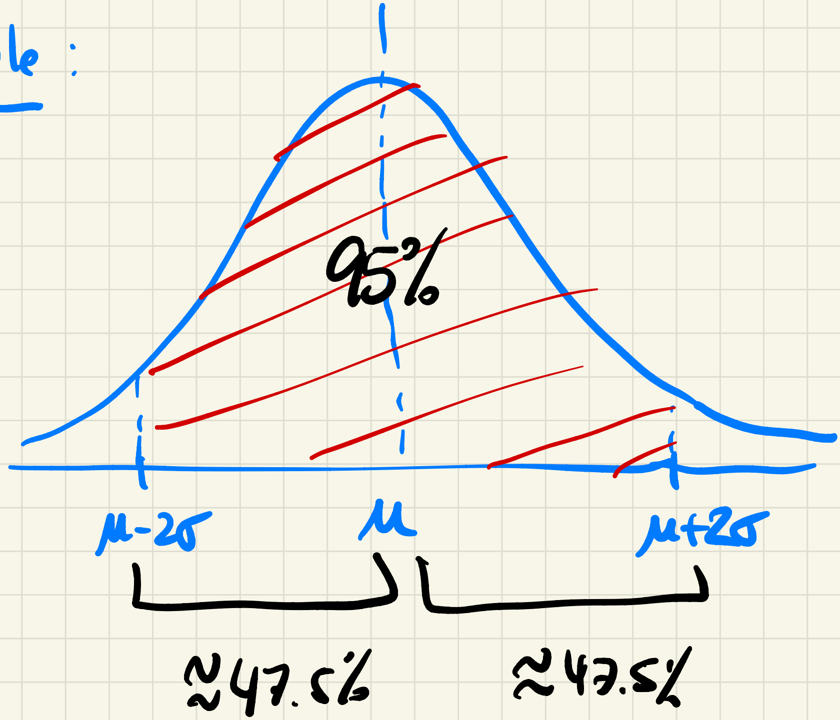
$X \sim N(\mu, \sigma)$

$$\begin{aligned} P(-b < X < b) &= P\left(-\frac{(b-\mu)}{\sigma} < \frac{X-\mu}{\sigma} < \frac{b-\mu}{\sigma}\right) \\ &= P\left(-\frac{(b-\mu)}{\sigma} < Z < \frac{b-\mu}{\sigma}\right) \end{aligned}$$

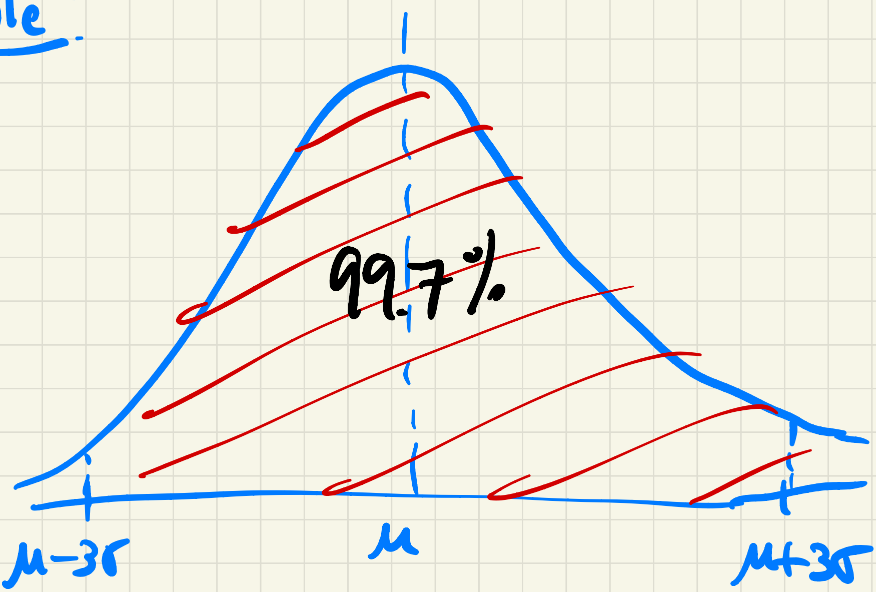
$Z \sim N(0,1)$

Note that $\frac{(b-\mu)}{\sigma}$ is called the Z-score of b. The Z-score of b is the number of standard deviations it is above or below the mean of the distribution. The fact that $P(-1 < Z < 1) = 0.683$ can be

2nd Rule:



3rd Rule:



interpreted as saying that the probability that a normal random variable falls within one standard deviation of its mean is 0.683. Therefore, in the statement "with high probability, normal random variables tend to be close to their mean value," *close* is interpreted as being relative to the standard deviation of the random variable.

3.10 Using Z-scores to Compare Top NFL Season Receiving Performance

The Z-score of a measurement Y is defined as

$$\frac{Y-\mu}{\sigma}$$

where μ is the mean of Y values for some distribution, and σ is the corresponding standard deviation. In the previous section, Z-scores were used as a way to understand, and calculate, probabilities regarding random variables with normal distributions. In this section, Z-scores are used to compare and standardize measurements.

The Z-score gives the number of standard deviations a measurement is above or below the mean. For example, if a measurement has a Z-score of 2, then that measurement is 2 standard deviations greater than the average value. Therefore, the Z-score takes into account both the average value of the measurement and the variability of the measurement, as measured by the standard deviation; Z-score give a simple way to compare a statistic for a particular player or team to the values obtained by other players or teams.

In 2012, Calvin Johnson had 1964 receiving yards, the highest yearly total up to that time. It is natural to ask how Johnson's 2012 season compares to other great years for receivers. In Table 3.3 shows the receiving yard totals for Johnson and 5 other receivers; these were chosen to represent a wide range of eras and are not necessarily the 5 best seasons in terms of receiving yards.

Table 3.3: Top receiving yard performances in different eras

Player	Year	Receiving Yards
Calvin Johnson	2012	1964
Marvin Harrison	2002	1722
Jerry Rice	1995	1848
John Jefferson	1980	1340
Otis Taylor	1971	1110
Raymond Berry	1960	1289

Direct comparison of the six seasons totals in Table 3.3 may be misleading because of the way the role of the passing game in the NFL has changed over the years. One way to account for these differences is to compare each receiver to the other receivers that played that season.

Using the Z-score approach, we convert each player's performance to a Z-score and then compare the Z-scores of the different players; the player with the highest Z-score has the best performance relative to his peers. Let Y_0 be the receiving yards for a player in a given year. For the mean and standard deviation of Y_0, we must use the sample-based values. Therefore, let $\overline{Y}_0$ denote the average receiving yards for some group of players and let S_0 denote the standard deviation of receiving yards for those players. Then, the Z-score of Y_0 is given by

$$\frac{Y_0 - \overline{Y}_0}{S_0}$$

To implement this approach, we need to choose players to use to calculate $\overline{Y}_0$ and S_0. One possibility is to use the set of all players catching at least one pass in the given year. Table 3.4 shows the average and standard deviation of receiving yards for those players for each year represented in Table 3.3. Table 3.4 also shows the Z-scores of the receiving yards for those players in Table 3.3.

Table 3.4: Mean and standard deviation of receiving yards for all players with at least one reception for the years represented in Table 3.3

Year	Average	SD	Z-score
2012	269.2	329.6	5.142
2002	291.2	325.8	4.392
1995	288.4	342.8	3.600
1980	280.5	278.4	3.806
1971	224.3	223.8	3.958
1960	234.0	263.9	4.032

$\frac{1964 - 269.2}{329.6}$

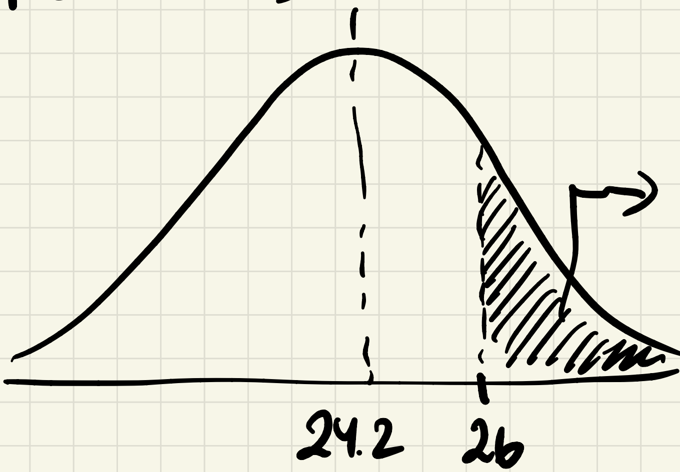
$\frac{1722 - 291.2}{325.8}$

According to the results of Table 3.4, Calvin Johnson's performance in 2012, which corresponds to a Z-score of about 5.1, is the most impressive, with this receiving yards over 5 standard deviations higher than the average receiving yards for players with at least 1 reception.

① X : number of points in a game

ⓐ $X \sim N(24.2, 2)$

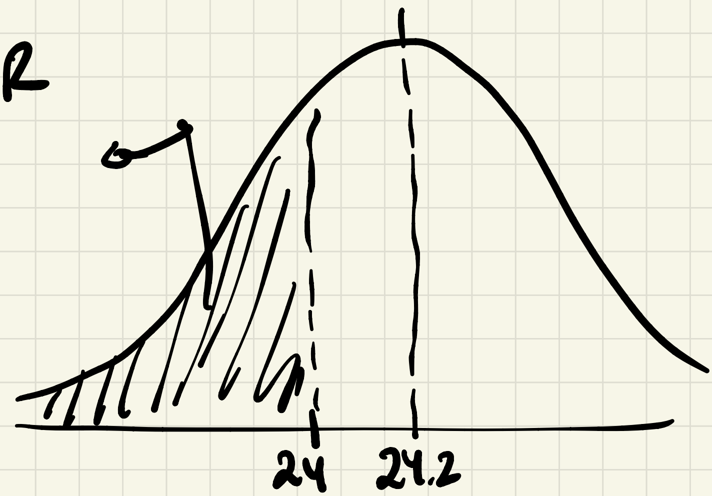
ⓑ $P(X > 26)$



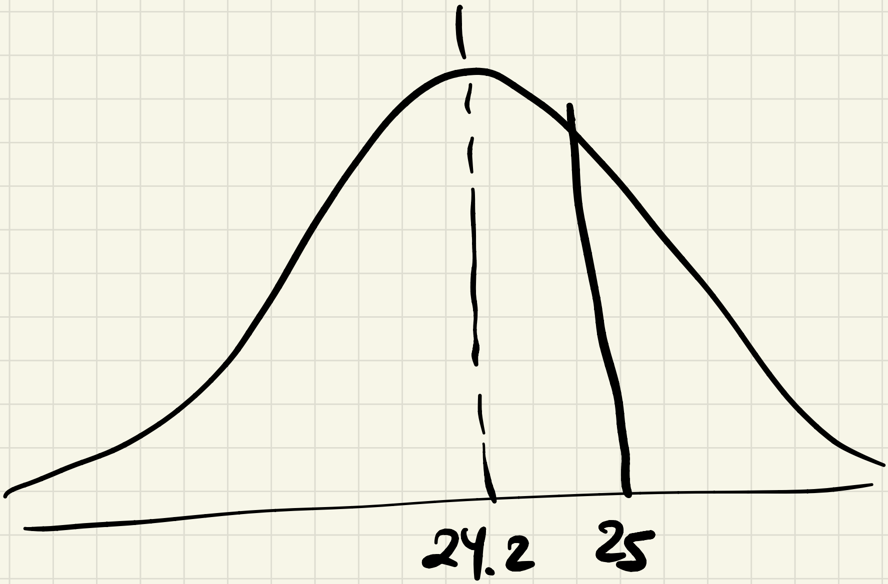
using R
18.4%

ⓒ $P(X < 24)$

using R
46%



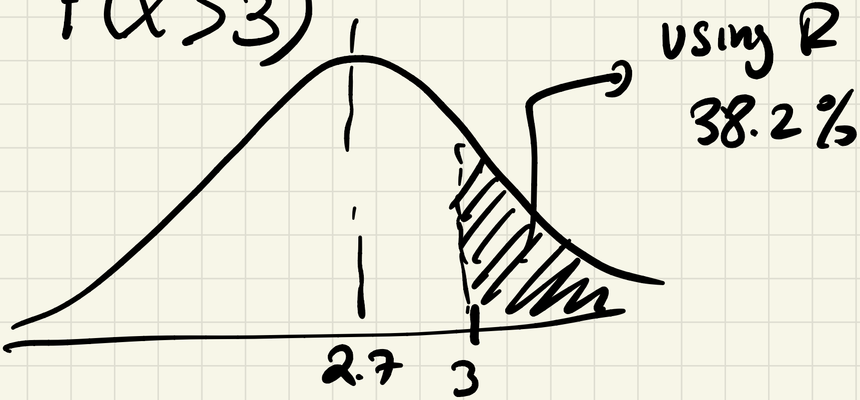
(d) $P(X=25)=0$



(2) X: number of touchdown passes per game

(a) $X \sim N(2.7, 1)$

(b) $P(X>3)$



(c) $P(X<2)$

Using R

24.2%