1 Probability Distributions

In the chapter about descriptive statistics samples were discussed, and tools introduced for describing the samples with numbers as well as with graphs.

In this chapter models for the population will be introduced. One will see how the properties of a population can be described in mathematical terms. Later we will see how samples can be used to draw conclusions about those properties. That step is called statistical inference.

Definition:
A variable $X$ (we use capital letters for random variables) is a random variable (rv) if the value that it assumes, corresponding to the outcome of an experiment, is a chance or random event.

Example:
- $X=$ number of observed "Tail" while tossing a coin 10 times
- $X=$ survival time after specific treatment of a randomly selected patient
- $X=$ SAT score for a randomly selected college applicant

Similar as for variables in sample data, rvs can be categorical or quantitative, and if they are quantitative they can be either discrete or continuous.

\[
\begin{align*}
\text{categorical} & \quad \uparrow \\
\text{random variable} & \quad \downarrow \\
\text{discrete} & \quad \uparrow \\
\text{quantitative} & \quad \downarrow \\
\text{continuous} & \quad \downarrow 
\end{align*}
\]

Similar to data description, the models for rvs depend entirely on the type the rv. The models for continuous rvs will be different than those for categorical or discrete rvs.

1.1 Categorical Random Variables

Categorical random variables are described by their distribution.

Definition
The probability distribution of a categorical rv is a table giving all possible categories the rv can assume and the associated probabilities.

Example:
The population investigated are the students of a selected college. The random variable of interest $X$ is the residence status, it can be either resident or nonresident.

If a student is chosen randomly from this college, the probability for being a resident is 0.73. Is $X$ the random variable resident status then write $P(X = \text{resident}) = 0.73$. 
The probability distribution is:

<table>
<thead>
<tr>
<th>resident status</th>
<th>probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>resident</td>
<td>0.73</td>
</tr>
<tr>
<td>nonresident</td>
<td>0.27</td>
</tr>
</tbody>
</table>

1.2 Numerical Random Variables

1.2.1 Discrete Random Variables

Remember: A discrete rv is one whose possible values are isolated points along the number line.

Definition:
The probability distribution for a discrete rv $X$ is a formula or table that gives the possible values of $X$, and the probability $p(X)$ associated with each value of $X$.

<table>
<thead>
<tr>
<th>Value of $X$</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>...</th>
<th>$x_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>$p_1$</td>
<td>$p_2$</td>
<td>$p_3$</td>
<td>...</td>
<td>$p_n$</td>
</tr>
</tbody>
</table>

The probabilities must satisfy two requirements:

- Every probability $p_i$ is a number between 0 and 1.
- $\sum p_i = 1$.

Example:
Toss two unbiased coins and let $x$ equal the number of heads observed.

The simple events of this experiment are:

<table>
<thead>
<tr>
<th>coin 1</th>
<th>coin 2</th>
<th>$X$</th>
<th>$p$(simple event)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>H</td>
<td>2</td>
<td>1/4</td>
</tr>
<tr>
<td>H</td>
<td>T</td>
<td>1</td>
<td>1/4</td>
</tr>
<tr>
<td>T</td>
<td>H</td>
<td>1</td>
<td>1/4</td>
</tr>
<tr>
<td>T</td>
<td>T</td>
<td>0</td>
<td>1/4</td>
</tr>
</tbody>
</table>

So that we get the following distribution for $x=$number of heads observed:

<table>
<thead>
<tr>
<th>$X$</th>
<th>$p(X)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1/4</td>
</tr>
<tr>
<td>1</td>
<td>1/2</td>
</tr>
<tr>
<td>2</td>
<td>1/4</td>
</tr>
</tbody>
</table>

With the help of this distribution we can calculate that $P(x \leq 1) = P(x = 0) + P(x = 1) = 1/4 + 1/2 = 3/4$.

Properties of discrete probability distributions:

- $0 \leq p(X) \leq 1$
The expected value or population mean $\mu$ (mu) of a rv $X$ is the value that you would expect to observe on average if the experiment is repeated over and over again. It is the center of the distribution.

**Definition:**
Let $X$ be a discrete rv with probability distribution $p(X)$. The *population mean* $\mu$ or *expected value* of $X$ is given as

$$
\mu = E(X) = \sum_x X p(X).
$$

**Example:**
The expected value of the distribution of $X=$the number of heads observed tossing two coins is calculated by

$$
\mu = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1
$$

**Definition:**
Let $X$ be a discrete rv with probability distribution $p(X)$. The *population variance* $\sigma^2$ of $X$ is

$$
\sigma^2 = E((X - \mu)^2) = \sum_x (X - \mu)^2 p(X).
$$

The *population standard deviation* $\sigma$ (sigma) of a rv $X$ is equal to the square root of its variance.

$$
\sigma = \sqrt{\sigma^2}
$$

**Example (continued):**
The population variance of $X=$number of heads observed tossing two coins is calculated by

$$
\sigma^2 = (0 - 1)^2 \cdot \frac{1}{4} + (1 - 1)^2 \cdot \frac{1}{2} + (2 - 1)^2 \cdot \frac{1}{4} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}
$$

and the population standard deviation is:

$$
\sigma = \sqrt{\sigma^2} = \frac{1}{\sqrt{2}}.
$$

### 1.2.2 Binomial Probability Distribution
Examples of discrete random variables can be found in a variety of everyday situations and across most academic disciplines. Here we will discuss one discrete probability distribution that serves as a model in lot of situations.

Many practical experiments result in data with only two possible outcomes

- patient survived 5 years (yes/no)
- coin toss (Head/Tail)
• poll: in favor of presidential decision (yes/no)
• person possesses a gene linked to Alzheimer (yes/no)

Each sampled person is the equivalent to tossing a coin, only that the probability for the event of interest does not have to be equal to 1/2.

**Definition:** A binomial experiment has the following four characteristics:

1. The experiment consists of \( n \) identical independent trials.
2. Each trial results in one of two possible outcomes (success \( S \), failure \( F \)).
3. The probability of success on a single trial is equal for all trials, \( p \). The probability of failure is then equal to \( q = 1 - p \).
4. We are interested in \( X \), the number of successes observed during \( n \) trials, for \( X = 0, 1, 2, \ldots, n \).

Before we discussed the example to toss a fair coin twice and determined the probability distribution for \( x = \) number of heads, this is a binomial distribution with \( n = 2 \) and \( p = 0.5 \). This can be done for any given \( n \) and \( p \), which are the parameter of a binomial distribution:

**The Binomial Probability Distribution**

A binomial experiment consists of \( n \) identical trials with probability of success \( p \) on each trial. The probability of \( k \) successes in \( n \) trials is (\( X \) is the random variable giving the number of successes)

\[
P(X = k) = \binom{n}{k} p^k q^{n-k}
\]

for values of \( k = 0, 1, 2, \ldots, n \).

With

\[
\binom{n}{k} = \frac{n!}{k!(n-k)!}
\]

**Theorem:**

Suppose \( X \) is a binomial distributed rv, with \( n \) trials, and probability of success \( p \). The population mean of \( X \) is

\[
\mu = E(X) = np
\]

The population variance of \( X \) is

\[
\sigma^2 = npq = np(1 - p)
\]

The population standard deviation of \( X \) is

\[
\sigma = \sqrt{npq} = \sqrt{np(1 - p)}
\]

**Example:** Assume an experiment with \( n = 10 \) and \( p = 0.1 \).

1. The mean \( \mu = 10(0.1) = 1 \)
2. The variance $\sigma^2 = 10(0.1)(1 - 0.1) = 0.9$

3. The standard deviation $\sigma = \sqrt{0.9} = 0.95$

4. The probability of 2 successes $P(x = 2) = \left(\frac{2}{10}\right)0.1^2 \cdot 0.9^8 = \frac{10!}{2!8!} \cdot 0.1^2 \cdot 0.9^8 = \frac{10 \cdot 9}{2} \cdot 0.01 \cdot 0.43 = 0.1937$

5. In order to give the complete probability distribution you would have to calculate $P(x = k)$ for $k = 0, 1, 2, \ldots, 10$

A probability histogram would look like this

Example:
It is known that a given marksman can hit a target on a single trial with probability equal to 0.8. Suppose he fires 4 shots at the target:

1. What is the probability of hitting at most one targets. Find for $n = 4$ and $p = 0.8$ the probability

$$P(x \leq 2) = P(X = 0) + P(X = 1) = \binom{4}{0}0.8^0 \cdot 0.2^4 + \binom{4}{1}0.8^1 \cdot 0.2^3 = 0.272$$

2. What is the probability to exactly hit three targets?

$$P(X = 3) = \binom{4}{3}0.8^3 \cdot 0.2 = 0.409$$

3. What is the probability to hit at least 1 target?

$$P(X \geq 1) = 1 - P(X = 0) = 1 - 0.002 = 0.998$$
1.2.3 Continuous Random Variables

Continuous data variables are described by histograms. For histograms the measurement scale is divided in class intervals and the area of the rectangles put above those intervals is proportional to the relative frequency of the data falling into this interval.

The relative frequency can be interpreted as an estimate of the probability for falling in the associated interval.

With this interpretation the histogram becomes an ”estimates” of the probability distribution of the continuous random variable.

Definition:
The probability distribution of a continuous random variable \( X \) is described by a density curve. The probability to fall within a certain interval is then given by the area under the curve above that interval.

1. The total area under a density curve is always equal to 1.

2. The area under the curve and above any particular interval equals the probability of observing a value in the corresponding interval when an individual or object is selected at random from the population.

We can calculate that the probability for falling in the interval \([-2; 0]\) equals 0.4772.

Example:
The density of a uniform distribution in an interval \([0; 5]\) looks like this:
Use the density function to calculate probabilities for a random variable $X$ with a uniform distribution on $[0; 5]$:

- $P(X \leq 3) = \text{area under the curve from } -\infty \text{ to } 3 = 3 \cdot 0.2 = 0.6$
- $P(1 \leq X \leq 2) = \text{area under the curve from } 1 \text{ to } 2 = 1 \cdot 0.2 = 0.2$
- $P(X > 3.5) = \text{area under the curve from } 3.5 \text{ to } \infty = 1.5 \cdot 0.2 = 0.3$

Remark: Since there is zero area under the curve above a single value, the definition implies for continuous random variables and numbers $a$ and $b$:

- $P(X = a) = 0$
- $P(X \leq a) = P(X < a)$
- $P(X \geq b) = P(x > b)$
- $P(a < X < b) = P(a \leq X \leq b)$

This is generally not true for discrete random variables.

How to choose a model for a given variable of a sample?
The model (density function) should resemble the histogram for the given variable. Fortunately, many continuous data variables have bell shaped histograms. The normal probability distribution provides a good model for modeling this type of data.

### 1.2.4 Normal Probability Distribution

The density function of a normal distribution is unimodal, mound shaped, and symmetric. There are many different normal distributions, they are distinguished from one another by their population mean $\mu$ and their population standard deviation $\sigma$. 
The **density function of a normal distribution** is given by

\[ f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)} \quad -\infty \leq x \leq \infty \]

with \( e \approx 2.7183 \) and \( \pi \approx 3.1416 \). \( \mu \) and \( \sigma \) are called the parameters of a normal distribution. And a normal distribution with mean \( \mu \) and standard deviation \( \sigma \) is denoted as \( N(\mu, \sigma) \).

\( \mu \) is the center of the distribution, right at the highest point of the density distribution function. At the values \( \mu - \sigma \) and \( \mu + \sigma \) the density curve has turning points. Coming from \(-\infty\) the curve turns from a left to a right curve at \( \mu - \sigma \) and again into in a left curve at \( \mu + \sigma \).
If the normal distribution is used as a model for a specific situation, the mean and the standard deviation have to be chosen for that situation. E.g. the height of students at a college follow a normal distribution with $\mu = 178$ cm and $\sigma = 10$ cm. The normal distribution is one example for a quantitative continuous distribution!

**Definition:**
The normal distribution with $\mu = 0$ and $\sigma = 1$ is called the *Standard Normal Distribution*, $\mathcal{N}(0,1)$.

In order to work with the normal distribution, we need to be able to calculate the following:

1. We must be able to use the normal distribution to compute probabilities, which are areas under the normal curve.

2. We must be able to describe extreme values in the distribution, such as the largest 5%, the smallest 1%, the most extreme 10% (which would include the largest 5% and the smallest 5%).

We first look how to compute these for a Standard Normal Distribution.

Since the normal distribution is a **continuous** distribution the following holds for every normal distributed random variable $X$:

$P(X < z) = P(X \leq z)$ area under the curve from $-\infty$ to $z$. 

The area under the curve of a normal distribute random variable is very hard to calculate. There is no simple formula that can be used to calculate the area.

Table II in Appendix A (in the text book) tabulates for many different values of $z$ the area under the curve from $-\infty$ to $z$ for standard normal distributed random variables. These values are called cumulative density function.

From now on use $Z$ to indicate a standard normal distributed random variable ($\mu = 0$ and $\sigma = 1$). Using the table you find that,

- $P(Z < 1.75) = P(0 < z \leq 1.75) = 0.9599$ therefore
- $P(Z > 1.75) = 1 - P(Z \leq 1.75) = 1 - 0.9599 = 0.0401$.
- $P(-1 < Z < 1) = P(Z < 1) - P(Z \leq -1) = .8413 - .1587 = .6826$. (Compare with the Empirical Rule.)
The shaded area equals 0.6826.

The first probability can be interpreted as meaning that, in a long sequence of observations from a Standard Normal distribution, about 95.99% of the observed values will fall below 1.75. Try this for different values!

Now we will look how to identify extreme values.

**Definition:**
For any particular number \( \alpha \) between 0 and 1, \( z_{\alpha} \) of a distribution is a value such that the cumulative area to the right of \( z_{\alpha} \) is \( \alpha \).

If \( X \) is a random variable then \( z_{\alpha} \) is given by:

\[
P(X \geq z_{\alpha}) = \alpha
\]

To determine \( z_{\alpha} \) for a standard normal distribution, we can use Table II in Appendix A again.
• Suppose we want to describe the values that make up the largest 2%. So we are looking for \( z_{0.02} \), then the area, which falls below \( z_{0.02} \) equals \( 1 - 0.02 = 0.98 \), therefore

\[
P(Z \leq z_{0.02}) = 0.98.
\]

So look in the body of the Table II for the cumulative area 0.98. The closest you will find 0.9798 for \( z_{0.02} = 2.05 \) This is the best approximation you can find from the table.

The result is that the largest 2% of the values of a standard normal distribution fall into the interval \([2.05, \infty)\).

• Suppose now we are interested in the smallest 5%. So we are looking for \( z^* \), with

\[
P(z < z^*) = 0.05 \iff P(z > z^*) = 1 - 0.05 = 0.95
\]

Therefore \( z^* = z_{0.95} \).

In the table we have to look for the area to the left, 0.05. Checking the Table we find values 0.0495 and 0.0505, with 0.05 exactly in the middle, so we take the average of the corresponding numbers and get

\[
z_{0.95} = \frac{(-1.64) + (-1.65)}{2} = -1.645
\]

• And now we are interested in the most extreme 5%. That means we are interested in the middle 95%. Since the normal distribution is the symmetric the most extreme 5% can be split up in the lower 2.5% and the upper 2.5%. Symmetry about 0 implies that \(-z_{0.025} = z_{0.975}\).
In Table II we find $z_{0.975} = -1.96$, so that $z_{0.025} = 1.96$.

We found the result, that the 5% most extreme values can be found outside the interval $[-1.96, 1.96]$.

Until here we covered how to find probabilities and values for $z_\alpha$ for the standard normal distribution. It remains to show how to find these for any normal distribution, hopefully using the results for the standard normal distribution.

**Lemma:** Is $X$ normal distributed with population mean $\mu$ and population standard deviation $\sigma$ then the standardized random variable

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

is normal distributed with $\mu = 0$ and $\sigma = 1$

or $Z \sim N(0, 1)$.

The following example illustrates how the probability and the percentiles can be calculated by using the standardization process from the Lemma.

**Example:** Let $X$ be normal distributed with $\mu = 100$ and $\sigma = 5$, $x \sim N(100, 5)$.

1. Calculate the area under the curve between 98 and 107 for the distribution chosen above.

$$P(98 < X < 107) = P\left(\frac{98-100}{5} < \frac{X-100}{\sigma} < \frac{107-100}{5}\right)$$

$$= P\left(-0.4 < Z < \frac{7}{5}\right)$$

$$= P(-0.4 < z < 1.4)$$

This can be calculated using Table II. $P(-0.4 < Z < 1.4) = P(Z < 1.4) - P(Z < -0.4) = (0.5 + 0.4192) - (0.5 - 0.1554) = 0.5746$.

2. To find $x_{0.97}$, the value so that the probability to fall above is 0.97, use

$$0.97 = P(X \geq x_{0.97})$$

$$0.03 = P(X \leq x_{0.97})$$

$$= P\left(\frac{X-100}{5} \leq \frac{x_{0.97} - 100}{5}\right)$$

But then $\frac{x_{0.97} - 100}{5}$ equals the $z_{0.97}$ from a standard normal distribution, which we can find in Table II (look up 0.03).

$$\frac{x_{0.97} - 100}{5} = -1.88$$

This is equivalent to $x_{0.97} = -1.88 \cdot 5 + 100 = 100 - 9.40 = 90.6$.

So that the lower 3% of a normal distributed random variable with mean $\mu = 100$ and $\sigma = 5$ fall into the interval $(-\infty, 90.6]$, or the top 97% from this distribution fall above 90.6.

**Example:**
Assume that the length of a human pregnancy follows a normal distribution with mean 266 days and standard deviation 16 days.
What is the probability that a human pregnancy lasts longer than 280 days?

\[
P(X > 280) = P\left( \frac{X - 266}{16} > \frac{280 - 266}{16} \right) \quad \text{standardize}
\]

\[
= P(Z > \frac{14}{16})
\]

\[
= 1 - P(Z \leq 0.875)
\]

\[
= 1 - 0.8078
\]

\[
= 0.1922
\]

How long do the 10% shortest pregnancies last? Find \(x_{0.9}\).

(Draw a picture.)

\[
P\left( \frac{X - 266}{16} \leq \frac{x_{0.9} - 266}{16} \right) = 0.1 \quad \text{standardize}
\]

\[
P(Z \leq \frac{x_{0.9} - 266}{16}) = 0.1
\]

So \(\frac{x_{0.9} - 266}{16} = z_{0.9}\).

\[
\frac{x_{0.9} - 266}{16} = z_{0.9} = -1.28 \quad \text{from Table II}
\]

This is equivalent to

\[
x_{0.9} = 16(-1.28) + 266 = 245.5 \text{ days}
\]

The 10% shortest pregnancies last shorter than 245.5 days.

### 1.3 Descriptive Methods for Assessing Normality

Process of assessing Normality

1. Construct a histogram or stem-and-leaf plot and note the shape of the graph. If the graph resembles a normal (bell-shaped) curve, the data could be normally distributed. Big deviations from a normal curve will make us decide that the data is not normal.

2. Compute the intervals \(\bar{x} \pm s\), \(\bar{x} \pm 2s\), and \(\bar{x} \pm 3s\). If the empirical rule holds approximately, the data might be normally distributed. If big deviations occur, the assumption might be unreasonable.

3. Construct a normal probability plot (normal Q-Q Plot) for the data. If the data are approximately normal, the points will fall (approximately) on a straight line.

Construction of normal probability plot:

1. Sort the data from smallest to largest and assign the rank, \(I\) (smallest gets 1, second 2, etc.)

2. Find the normal scores for the relative rank (Table III).

3. Construct a scatter plot of the measurements and and the expected normal scores.

**Example 1**

Assume we have the data set 1, 2, 2.5, 3, 9

This is not enough data to do a histogram, so that method does not work. Let’s do a normal probability plot to investigate.
The first 4 points seem to fall on a line, but the fifth does not fit and seems to be an outlier, or uncommon value. The data do not seem to come from a normal distribution. Outliers are highly unlikely in normal distributions.
2 Binomial Probability Distribution

Examples of discrete random variables can be found in a variety of everyday situations and across most academic disciplines. Here we will discuss one discrete probability distribution that serves as a model in lot of situations.

Many practical experiments result in data with only two possible outcomes

- patient survived 5 years (yes/no)
- coin toss (Head/Tail)
- poll: in favor of presidential decision (yes/no)
- person possesses a gene linked to Alzheimer (yes/no)

Each sampled person is the equivalent to tossing a coin, only that the probability for the event of interest does not have to be equal to 1/2.

**Definition:** A binomial experiment has the following four characteristics:

1. The experiment consists of \( n \) identical independent trials.
2. Each trial results in one of two possible outcomes (success \( S \), failure \( F \)).
3. The probability of success on a single trial is equal for all trials, \( p \). The probability of failure is then equal to \( q = 1 - p \).
4. We are interested in \( X \), the number of successes observed during \( n \) trials.

Before we discussed the example to toss a fair coin twice and determined the probability distribution for \( X = \) number of heads, this is a binomial distribution with \( n = 2 \) and \( p = 0.5 \).

\( n \), the number of trials, and \( p \), the probability for Success, are the parameters of a binomial distribution:

**The Binomial Probability Distribution**

A binomial experiment consists of \( n \) identical trials with probability for success \( p \) for each trial. The probability for \( k \) successes in \( n \) trials is then

\[
P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}
\]

for values of \( k = 0, 1, 2, \ldots, n \), where

\[
\binom{n}{k} = \frac{n!}{k!(n-k)!}, \quad \text{and} \quad n! = n(n-1)(n-2)\cdots2\cdot1
\]

**Theorem:**

Suppose \( X \) is a binomial distributed rv, with \( n \) trials, and probability of success \( p \). The population mean of \( X \) is

\[
\mu = E(X) = np
\]
The population variance of $X$ is

$$\sigma^2 = npq = np(1 - p)$$

The population standard deviation of $x$ is

$$\sigma = \sqrt{npq} = \sqrt{np(1 - p)}$$

**Example:** Assume an experiment with $n = 10$ and $p = 0.1$.

1. The mean $\mu = 10(0.1) = 1$
2. The variance $\sigma^2 = 10(0.1)(1 - 0.1) = 0.9$
3. The standard deviation $\sigma = \sqrt{0.9} = 0.95$
4. The probability of 2 successes $P(x = 2) = C_{10}^2 \cdot 0.1^2 \cdot 0.9^8 = \frac{10!}{2!8!} \cdot 0.1^2 \cdot 0.9^8 = \frac{10 \cdot 9}{2} \cdot 0.01 \cdot 0.43 = 0.1937$
5. In order to give the complete probability distribution you would have to calculate $P(x = k)$ for $k = 0, 1, 2, \ldots, 10$

A probability histogram would look like this

![Probability histogram](image)

**Example:**
It is known that a given marksman can hit a target on a single trial with probability equal to 0.8. Suppose he fires 4 shots at the target:
1. What is the probability of hitting at most two targets. Find for \( n = 4 \) and \( p = 0.8 \) the probability.

\[
P(x \leq 2) = P(x = 0) + P(x = 1) + P(x = 2) = 0.2^4 + \binom{4}{1}0.8^10.2^3 + \binom{4}{2}0.8^20.2^2 = 0.1808
\]

2. What is the probability to exactly hit three targets?

\[
P(x = 3) = \binom{4}{3}0.8^30.2^1 = 0.4096
\]

3. What is the probability to hit at least 1 target?

\[
P(x \geq 1) = 1 - P(x < 1) = 1 - P(x = 0) = 1 - \binom{4}{0}0.8^00.2^4 = 0.9984
\]

Instead of using the formula for the binomial distribution we also can find the approximate probabilities for the binomial distribution with the help of the normal distribution. This only works if the binomial distribution is symmetric enough. If it is we find an appropriate z-score, and read the probability from the normal table.

1. Find \( n \) and \( p \) for the problem and calculate the interval

\[
\mu - 3\sigma, \mu + 3\sigma
\]

If the interval falls between 0 and \( n \), this method is appropriate.

2. Express the probability you want to find using terms like \( P(X \leq a) \)

3. Use the continuity correction \( a + 0.5 \) and standardize

\[
z = \frac{(a + 0.5) - \mu}{\sigma}
\]

4. Find \( P(Z \leq z) \) from table IV to approximate \( P(X \leq a) \).

Let’s do an example

**Example 2**

The probability to get an A or a B in Stat 151 is 0.35. In a class of 60, what is the probability that less than 15 people get an A.

Go through the 4 steps

1. \( n = 60, p = 0.35, \mu = n(p) = 21 \sigma = \sqrt{n(p)(1-p)} = 3.695 \), then

\[
\mu - 3\sigma, \mu + 3\sigma \rightarrow 9.916; 32.085
\]

This interval falls between 0 and 60, we can use this method.

2. We want to find \( P(X < 15) = P(X \leq 14) \) \( (a = 14) \)

3. 

\[
z = \frac{15.5 - 21}{3.695} = -1.488
\]

4. \( P(X \leq 14) \approx P(Z \leq -1.488) = 0.5 - P(0 \leq Z \leq 1.49) = 0.5 - 0.4319 = 0.0681 \)

The probability that less than 14 students out of 60 will get an A or B in Stat 151 is approximately 0.0681.