Review of Probability

Probability

1. We have discrete and continuous probability distributions.
2. For continuous probability distributions we can show a curve of the probability density function.
3. The area under the entire curve is 1.00.
4. The area under the curve and to the left of a value **v** is the probability that the random variable **X** is less than **v**. This is written as **P(X<v)**.
5. For a continuous probability, since there is no "area" under the curve at a value **v**, **P(X=v) = 0.** Therefore, **P(X < v) = P(X ≤ v)**.
6. For symmetric probability distributions we have **P(X < v) = P(X > -v)**.
7. For two values **a** and **b**, where **a < b**, we have
   P(a < X < b) = 1 - P(X<a or X>b).

The Normal Distribution

8. The **normal** distribution is a continuous probability distribution.
9. The **normal** distribution is based on a mathematical formula.
10. The **standard normal distribution** has **mean=0** (**μ=0**) and **standard deviation=1** (**σ=1**).
11. We have a **table** to find the area under the **standard normal** distribution probability density function curve and to the left of a value **z**, i.e., **P(X<z)**.
12. We have a function in R called **pnorm()** to find the area under the **standard normal** density function to the left of a value **z**, **pnorm(z)**.
13. We know how to read the table "backwards" so that if we are given a probability **p** we can find the value **z** that makes **P(X<z)=p** true.
14. We have a function in R called **qnorm()** that finds the value **z** needed to make the **P(X<z)** be the value of a specified probability **p**, that is **qnorm(p)** produces the value **z** so that **P(X<z)=p**.
15. We use **N(μ, σ)** to symbolize a distribution that is **normal** with **mean=μ** and **standard deviation=σ**.
16. The standard normal distribution is **N(0,1)**.

$$z = \frac{x - \mu}{\sigma}$$

17. We use   to map a non-standard normal distribution, **N(μ,σ)** to **N(0,1)**.
18. Both **pnorm()** and **qnorm()** have extended forms that allow us to use them directly with non-standard normal distributions.

The Sample Mean and Its Mean and Standard Deviation

Our task here is to look at the **mean** and **standard deviation** of of samples in general. We start with some **original population** that has a known **mean, μ,** and **standard deviation, σ.** We take repeated

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

samples, each of size **n**, and for each sample we generate a **sample mean,** , (and,
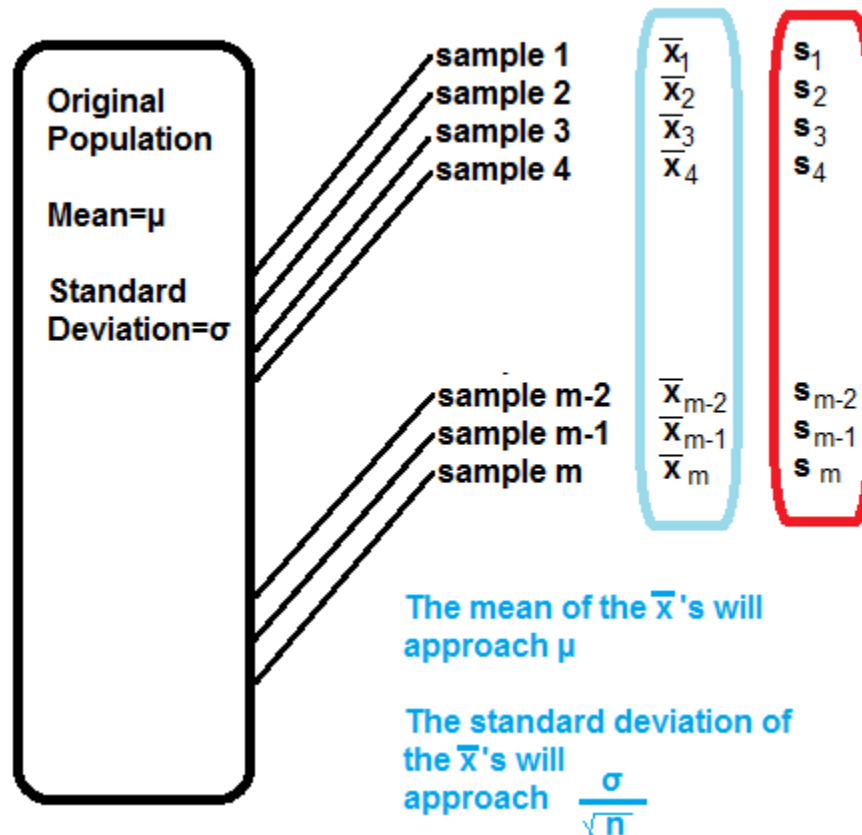
$$s = \sqrt{\frac{\sum_{i=1}^{n} \left(x_i - \overline{x}\right)^2}{n-1}}$$

although we will not make much use of it, the **sample standard deviation,**
). Then, we want to treat the collection of **sample means** as a population itself and we want to examine the **mean** and **standard deviation** of that population. Figure 1 attempts to show this.

Figure 1

Take repeated samples, each of size n, with replacement

As suggested in the above figure, the **mean** of all the **sample mean values** should be really close to the **mean of the original population**. Furthermore, the **standard deviation of the sample means** should be really close to the value $\dfrac{\sigma}{\sqrt{n}}$.

Confidence Intervals, based on the sample mean, σ known

The process of computing a **confidence interval** in the case where we **know the population standard deviation** and where we have a sample of size **n** that yields a **sample mean** $\overline{x}$ is as follows:

1. From the **confidence level** compute the value of $\dfrac{\alpha}{2}$ using $\dfrac{\alpha}{2} = \dfrac{1 - cl}{2}$

2. Use **qnorm(** $\frac{\alpha}{2}$ **)** to find the associated **z-score**, $z_{\frac{\alpha}{2}}$

3. Find the **margin of error** as $\left| z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right|$ parts to the **confidence interval** by evaluating

$$= \overline{x} \pm z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

We should be able to describe these same actions using R statements inside an R function. Consider the following function definition:

```
ci_known <- function( sigma=1, n=30, x_bar=0, cl=0.95)
  {
    # try to avoid some common errors
    if( cl <=0 | cl>=1)
      {return("Confidence interval must be strictly between 0.0 and 1")
      }
    if( sigma < 0 )
      {return("Population standard deviation must be positive")}
    if( n <= 1 )
      {return("Sample size needs to be more than 1")}
    if( as.integer(n) != n )
      {return("Sample size must be a whole number")}
    # to get here we have some "reasonable" values
    samp_sd <- sigma/sqrt(n)
    z <- abs( qnorm( (1-cl)/2))
    moe <- z*samp_sd
    low_end <- x_bar - moe
    high_end <- x_bar + moe
    result <- c(low_end, high_end, moe, samp_sd)
    names(result)<-c("CI Low","CI High","MOE", "Std Error")
    return( result )
  }
```

Confidence Intervals, based on the sample mean, σ unknown

The concept of a **confidence interval** was developed in the previous section. There we saw that we can take the **sample mean,** $\overline{x}$, as a **point estimate** of a **population mean**, and form an interval with the **point estimate** as the center of the interval and the lower and upper ends of the interval set at a distance, called the **margin of error**, on either side of the **point estimate**. Furthermore, we used the fact that the **distribution of sample means** is **normal** to help us determine the value of the **margin of error**.

Part of that determination was to state the **level of confidence** that we want in the process. Thus, we create a **95% confidence interval** by a process that will generate **confidence intervals** wide enough so that **95%** of the **confidence intervals** generated by this process will contain the true population mean, **μ**.

To do this we used that the fact that the **distribution of the sample mean** is not only **normal** but it has a

mean of **μ** and a **standard deviation =** $\frac{\sigma}{\sqrt{n}}$. That meant that for a **95% confidence interval** we needed

to find a **z-score**, $z_{\frac{\alpha}{2}}$, such that **95%** of the area under the **standard normal curve** is between the **z-score**

and its negative. Once we have the **z-score** then the **margin of error** is just $\left| z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right|$ and the

confidence $= \bar{x} \pm z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$ **interval**.

That whole process depends upon us knowing the value of **σ**, the **population standard deviation**The whole process is designed to find a **confidence interval** for the **population mean**. It would be absurd to do this if we already know the **population mean**. The process above requires that we know the **population standard deviation**. Since the **population standard deviation** is the **root mean squared deviation from the mean**, we need to know the **population mean** in order to compute the **population standard deviation**. It does raise the question of why we developed a process that we will never use?

We introduce the process above because it is a great model for creating a **confidence interval** when the population standard deviation is **unknown**. In such a situation we use the **sample standard deviation, s_x,** as a approximation to the **population standard deviation**. When we do this can no longer compute the

standard deviation of the **sample mean** by $\frac{\sigma}{\sqrt{n}}$ because we do not know **σ**. Instead, we compute the

standard deviation of the sample mean by $\frac{s_x}{\sqrt{n}}$. However, computed this way, we cannot say that the **sample mean** is **N(μ,s_x)**, instead we say that the **sample mean has a Student's t distribution with n-1 degrees of freedom.** Knowing this we can just mimic the earlier process, substituting the use of the

Student's t with the appropriate **degrees of freedom** for our $z_{\frac{\alpha}{2}}$ value.

The process of computing a **confidence interval** in the case where we **do not know the population standard deviation** and where we have a sample of size **n** that yields a **sample mean** $\bar{x}$ and a **sample standard deviation** s_x is as follows:

1. From the **confidence level** compute the value of $\frac{\alpha}{2}$ using $\frac{\alpha}{2} = \frac{1 - cl}{2}$

2. Use **qt(** $\frac{\alpha}{2}$ **)** to find the associated **t-score**, $t_{\alpha/2}$
3. Find the **margin of error** as

$$= \left| t_{\frac{\alpha}{2}} \cdot \frac{s_x}{\sqrt{n}} \right|$$

4. Find the two parts to the **confidence interval** by evaluating

$$\bar{x} \pm t_{\frac{\alpha}{2}} \cdot \frac{s_x}{\sqrt{n}}$$

We should be able to describe these same actions using R statements inside an R function. Consider the following function definition:

```
ci_unknown <- function( s=1, n=30,
                        x_bar=0, cl=0.95)
{
  # try to avoid some common errors
  if( cl <=0 | cl>=1)
  {return("Confidence interval must be strictly between 0.0 and 1")
  }
  if( s < 0 )
  {return("Sample standard deviation must be positive")}
  if( n <= 1 )
  {return("Sample size needs to be more than 1")}
  if( as.integer(n) != n )
  {return("Sample size must be a whole number")}
  # to get here we have some "reasonable" values
  samp_sd <- s/sqrt(n)
  t <- abs( qt( (1-cl)/2, n-1))
  moe <- t*samp_sd
  low_end <- x_bar - moe
  high_end <- x_bar + moe
  result <- c(low_end, high_end, moe, samp_sd)
  names(result)<-c("CI Low","CI High",
                   "MOE", "Std Error")
  return( result )
}
```

Confidence Intervals for Proportions

In the case where we have a population where some of the elements in that population have a specific characteristic, we talk about the **proportion of the population that has that characteristic**. We generally signify that **proportion** as **p**. If we take a sample of size **n** of that population and in that sample we find **x** items with the characteristic, then the value $\hat{p} = \frac{x}{n}$ is a **point estimate for p**. We want a **confidence interval** for **p** derived from the **sample**. That **confidence interval** will be

point estimate ± (margin of error)
$\hat{p} \pm$ **(margin of error)**

The we recall that for cases where **n*p≥10** and **n*(1-p)≥10** we have the distribution of $\widehat{p}$ is **normal** with **mean = p** and **standard deviation = $\sqrt{p*(1-p)/n}$**. Therefore, for a specified

$$z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{p \cdot (1-p)}{n}}$$

confidence level we can find $z_{\alpha/2}$ to make the **margin of error =**   That makes the **confidence interval** be

$$\widehat{p} \pm z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{p \cdot (1-p)}{n}}$$

Of course, the problem with this is that we do not know the value of **p** so we cannot compute that formula.

Instead, if we change our special conditions we can use $\widehat{p}$ instead of **p** in the approximation of the **standard deviation**. The new restrictions are

- The sample size **n** is no more than 5% of the population size. (Another, popular, way to say this is that the population size is more than 20 times the sample size, **n**.)
- Population items either have the characteristic or the do not. That is another way of saying that the population items fall into one of two categories, those with the characteristic and those without the characteristic.
- The sample must contain at least 10 items in each of the two categories.

$$\sqrt{\frac{\widehat{p} \cdot (1-\widehat{p})}{n}}$$

If these conditions are met then we can use   as the approximation of the **standard deviation** of the $\widehat{p}$'s. That gives the formula

$$\widehat{p} \pm z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\widehat{p} \cdot (1-\widehat{p})}{n}}$$

to use to find the desired **confidence interval** for the population proportion **p** based on a **sample** of size **n** with **x** items in the sample having the characteristic so that $\widehat{p}$ **=x/n**.

Confidence Intervals: Standard Deviation

On this section we look at creating a **confidence interval** for the **population standard deviation** or **variance** based on a **sample standard deviation**. We will do this for cases where we have a strong belief that the underlying population is approximately **normal**.

For populations that are approximately **normal** and for samples of size **n** the ratio of the quantity **(n-1)s²** to **σ²** is distributed as a **χ²** with **n-1** degrees of freedom. The consequence of this is that for a sample of size **n** with a **sample standard deviation of s_x** we can find a 95% confidence interval by computing the two

values $\sqrt{\dfrac{(n-1)s_x^2}{\chi^2_{R0.025}}}$ and $\sqrt{\dfrac{(n-1)s_x^2}{\chi^2_{L0.025}}}$, both for a **χ²** with **n-1** degrees of freedom and where we

understand that $\chi^2_{R0.025}$ represents the **x-value** that has an area equal to 0.025 to its **right** and $\chi^2_{L0.025}$ represents the **x-value** that has an area equal to 0.025 to its **left**. *You might note a seeming inversion here.*

The lower value for the confidence interval, $\sqrt{\dfrac{(n-1)s_x^2}{\chi^2_{R0.025}}}$ *, uses the area on the right, while the upper value*

of the confidence interval, $\sqrt{\dfrac{(n-1)s_x^2}{\chi^2_{L0.025}}}$ *, uses the area on the left. This inversion is the result of having the χ² critical values in the denominator. We know that the χ² value on the left will be a smaller value than the one on the right. Therefore, when we divide the numerator by those values the division by the larger denominator produces a smaller quotient.*

Hypothesis testing, based on the sample mean, σ known

The situation is:

1. We have a hypothesis about the "true" value of a population mean. That is, someone (perhaps us) claims that **H₀: μ = a**, for some value **a**.
2. By some miracle, we happen to know the **population standard deviation, σ**.
3. We will consider an alternative hypothesis which is one of the following
   - **H₁: μ > a**,
   - **H₁: μ < a**, or
   - **H₁: μ ≠ a**.
4. We want to test **H₀ against H₁**.
5. We have already determined the **level of significance** that we will use for this test. The **level of significance, α**, is the chance that we are willing to take that we will make a **Type I error**, that is, that we will **reject H₀ when, in fact, it is true**.

> Immediately, we recognize that samples of size **n** drawn from this population with have a distribution of the **sample mean** that is a **normal with mean=μ** and **standard deviation=σ/sqrt(n)**. At this point we proceed via the **critical value approach** or by the **attained significance approach.** These are just different ways to create a situation where we can finally make a decision. The **critical value approach** tended to be used more often when everyone needed to use the tables. The **attained significance approach** is more commonly used now that we have calculators and computers to do the computations. Of course either approach can be done with tables, calculators, or computers. Either approach gives the same final result.

Critical Value Approach

---

1. We find the **z-score** that corresponds to having the **level of significance** area more extreme than that **z-score**, remembering that if we are looking at being either too low or too high then we need half the area in both extremes.

2. We determine a sample size **n**.
3. Compute the $s_x = \sigma / sqrt(n)$ and use that value to compute $(z)(s_x)$.
4. Set the **critical value** (or values in the case of a two-sided test) such that it (they) mark the value(s) that is (are) that distance, $(z)(s_x)$, away from the mean given by **H_0: μ = a**.
5. Then, we take a random sample of size **n** from the population.
6. We compute the **sample mean, $\overline{x}$**.
7. If that mean $\overline{x}$ is more extreme than the **critical value(s)** then we say that **"we reject H_0 in favor of the alternate H_1"**. If the **sample mean** is not more extreme than the **critical value(s)** then we say **"we have insufficient evidence to reject H_0"**.

Attained significance Approach

1. Then, we take a random sample of size **n** from the population.
2. We compute the **sample mean, $\overline{x}$**.
3. Assuming that **H_0** is true, we compute the probability of getting the value $\overline{x}$ or a value more extreme than that. We can do this using the fact that the distribution of sample means for samples of size **n** is **normal** with **mean=μ** and **standard deviation=σ/sqrt(n)**.
4. If the resulting probability is **smaller** than or **equal to** the predetermined **level of significance** then we say that **"we reject H_0 in favor of the alternate H_1"**. If the resulting probability is not less than the predetermined **level of significance** then we say **"we have insufficient evidence to reject H_0"**.

Hypothesis testing, based on the sample mean, σ unknown

The situation is:

1. We have a hypothesis about the "true" value of a population mean. That is, someone (perhaps us) claims that **H_0: μ = a**, for some value **a**.
2. Not unexpectedly, we do not know the **population standard deviation, σ**.
3. We will consider an alternative hypothesis which is one of the following
    o   **H_1: μ > a**,
    o   **H_1: μ < a**, or
    o   **H_1: μ ≠ a**.
4. We want to test **H_0 against H_1**.
5. We have already determined the **level of significance** that we will use for this test. The **level of significance, α**, is the chance that we are willing to take that we will make a **Type I error**, that is, that we will **reject H_0 when, in fact, it is true**.

6.

Immediately, we recognize that samples of size **n** drawn from this population with have a distribution of the **sample mean** that is a **Student's t** with **n-1** degrees of freedom. At this point we proceed via the

> **critical value approach** or by the **attained significance approach.** These are just different ways to create a situation where we can finally make a decision. The **critical value approach** tended to be used more often when everyone needed to use the tables. The **attained significance approach** is more commonly used now that we have calculators and computers to do the computations. Of course either approach can be done with tables, calculators, or computers. Either approach gives the same final result.

Critical Value Approach

1. We determine a sample size **n**. This will set the number of degrees of freedom as we use the **Student's t distribution**.
2. We find the **t-score** that corresponds to having the **level of significance** area more extreme than that **t-score**, remembering that if we are looking at being either too low or too high then we need half the area in both extremes.
3. Then, we take a random sample of size **n** from the population.
4. We compute the **sample mean, $\overline{x}$** and the **sample standard deviation, s**
5. Compute $s_x = s\, / \,sqrt(n)$ and use that value to compute $(t)(s_x)$.
6. Set the **critical value** (or values in the case of a two-sided test) such that it (they) mark the value(s) that is (are) that distance, $(t)(s_x)$, away from the mean given by **H_0: μ = a**.
7. If the **sample mean**, $\overline{x}$, is more extreme than the **critical value(s)** then we say that **"we reject H_0 in favor of the alternate H_1"**. If the **sample mean** is not more extreme than

Attained significance Approach

1. Then, we take a random sample of size **n** from the population.
2. We compute the **sample mean, $\overline{x}$** and the **samle standard deviation, s**.
3. Assuming that **H_0** is true, we compute the probability of getting the value $\overline{x}$ or a value more extreme than the sample mean. We can do this using the fact that the distribution of sample means for samples of size **n** is a **Student's t** distribution with **n-1 degrees of freedom** with **mean=μ** and **standard deviation=s/sqrt(n)**.
4. If the resulting probability is **smaller** than or **equal to** the predetermined **level of significance** then we say that **"we reject H_0 in favor of the alternate H_1"**. If the resulting probability is not less than the predetermined **level of significance** then we say **"we have insufficient evidence to reject H_0"**.

the **critical value(s)** then we say **"we have insufficient evidence to reject H₀"**.

Hypothesis testing for population proportion

The situation is:

1. We have a population the members of which fall into two groups, those with a particular characteristic and those without that characteristic. *(As a small aside, the population may fall into many groups but we focus on one characteristic and lump the rest into a group of items not having that characcteristic.)*
2. We are interested in the proportion, **p**, of the population with the particular characteristic.
3. We have a hypothesis about the "true" value of the population **proportion**. That is, someone (perhaps us) claims that **H₀: p = a**, for some value **a**.
4. We will consider an alternative hypothesis which is one of the following
   - **H₁: p > a**,
   - **H₁: p < a**, or
   - **H₁: p ≠ a**.
5. We want to test **H₀ against H₁**.
6. We have already determined the **level of significance** that we will use for this test. The **level of significance, α**, is the chance that we are willing to take that we will make a **Type I error**, that is, that we will **reject H₀ when, in fact, it is true**.

7.

> Immediately, we recognize that samples of size **n** drawn from this population with have a distribution of the **sample proportion** that is a **normal** with **mean=p** and **standard deviation=sqrt(p*(1-p)/n)**. At this point we proceed via the **critical value approach** or by the **attained significance approach.** These are just different ways to create a situation where we can finally make a decision. Of the two, the **attained significance approach** is more commonly used. Either approach gives the same final result.

Critical Value Approach	Attained significance Approach
1. Using the **normal distribution** we find the **z-score** that corresponds to having the **level of significance** area more extreme than that **z-**	1. We determine a sample size **n**. In doing this we need to be sure that both **(n)(p)≥10** and **(n)(1-p)≥10**.

score, remembering that if we are looking at being either too low or too high then we need half the area in both extremes.

2. We determine a sample size **n**. In doing this we need to be sure that both **(n)(p)≥10** and **(n)(1-p)≥10**.
3. Also, we cannot sample more than 5% of the population. That means that the size of the population must be more than 20 times **n**.

$$\sigma_{\widehat{p}} = \sqrt{\frac{p \cdot (1-p)}{n}}$$

4. Compute and use that value to compute $z \cdot \sigma_{\widehat{p}}$.
5. Set the **critical value** (or values in the case of a two-sided test) such that it (they) mark the value(s) that is (are) that distance, $z \cdot \sigma_{\widehat{p}}$, away from the **proportion** given by **H₀: p = a**.
6. Then, we take a random sample of size **n** from the population.
7. We compute the **sample proportion,** $\widehat{p}$.
8. If that proportion $\widehat{p}$ is more extreme than the **critical value(s)** then we say that **"we reject H₀ in favor of the alternate H₁"**. If the **sample proportion** is not more extreme than the **critical value(s)** then we say **"we have insufficient evidence to reject H₀"**.

2. Also, we cannot sample more than 5% of the population. That means that the size of the population must be more than 20 times **n**.
3. Then, we take a random sample of size **n** from the population.
4. We compute the **sample proportion,** $\widehat{p}$.

$$\sigma_{\widehat{p}} = \sqrt{\frac{p \cdot (1-p)}{n}}$$

5. Compute .

$$z = \frac{\widehat{p} - p}{\sigma_{\widehat{p}}}$$

6. Compute .
7. Using the **standard normal distribution**, and taking into account the alternative hyposthesis, H₁, so that we know if we are doing a one-tail or two-tail test, we compute the probability of getting the value **z** or a value more extreme than that.
8. If the resulting probability is **smaller** than or **equal to** the predetermined **level of significance** then we say that **"we reject H₀ in favor of the alternate H₁"**. If the resulting probability is not less than the predetermined **level of significance** then we say **"we have insufficient evidence to reject H₀"**.

Hypothesis testing for normal population standard deviation

The situation is:

1. We have a population with a parameter that is **normally distributed**. For this situation it is important that the parameter has a normal distribution but we do not need to know, ahead of time, the mean or standard deviation of that distribution.
2. We are interested in the standard deviation, σ, of the population.
3. We have a hypothesis about the "true" value of the population **standard deviation**. That is, someone (perhaps us) claims that **H$_0$: σ = a**, for some value **a**.
4. We will consider an alternative hypothesis which is one of the following
   - **H$_1$: σ > a**,
   - **H$_1$: σ < a**, or
   - **H$_1$: $\sigma \neq$ a**.
5. We want to test **H$_0$ against H$_1$**.
6. We have already determined the **level of significance** that we will use for this test. The **level of significance, α**, is the chance that we are willing to take that we will make a **Type I error**, that is, that we will **reject H$_0$ when, in fact, it is true**.

> Immediately, we recognize that samples of size **n** drawn from this population with have a distribution of the **ratio** of **n-1** times the **sample variance** to the **population variance** that is a χ^2 with **n-1** degrees of freedom. Thus, if **H$_0$** is true and the population standard deviation is **a**,
>
> $$\chi^2 = \frac{(n-1) \cdot s^2}{a^2}$$
>
> then for samples of size **n** the statistic   will have a χ^2 distribution with **n-1** degrees of freedom.
>
> At this point we proceed via the **critical value approach** or by the **attained significance approach.** These are just different ways to create a situation where we can finally make a decision. The **critical value approach** tended to be used more often when everyone needed to use the tables. The **attained significance approach** is more commonly used now that we have calculators and computers to do the computations. Of course either approach can be done with tables, calculators, or computers. Either approach gives the same final result.

7. We determine a sample size **n**. This will set the number of degrees of freedom as we use the χ^2 distribution.

Critical Value Approach

1. Find the value or values (depending upon if we are looking at a one-tailed or a two-tailed test) of χ^2 that have the level of significance area

Attained significance Approach

1. We compute the **sample standard deviation, σ**.

outside of the value or values. For a two-tailed test (H_1: $\sigma \neq a$) we need to split the level of significance to that half is at the low end and half is at the right end. *[This is made somewhat more complex if you are using a table since most tables only give the upper tail value and because χ^2 is not a symmetric distribution.]* Use that value or those values as the **critical values**.

2. Take the random sample of size **n** and compute the sample standard deviation, **s**, from that sample.

3. Compute the value

$$\chi^2 = \frac{(n-1) \cdot s^2}{a^2}$$
.

4. If that value is beyond the **critical value(s)** then we say that **"we reject H_0 in favor of the alternate H_1"**. If that value is not more extreme than the **critical value(s)** then we say **"we have insufficient evidence to reject H_0"**.

$$\chi^2 = \frac{(n-1) \cdot s^2}{a^2}$$

2. Compute   .

3. Using the χ^2 **distribution**, and taking into account the alternative hyposthesis, H_1, so that we know if we are doing a one-tail or two-tail test, we compute the probability of getting the value **x2** or a value more extreme than that.

4. If the resulting probability is **smaller** than or **equal to** the predetermined **level of significance** then we say that **"we reject H_0 in favor of the alternate H_1"**. If the resulting probability is not less than the predetermined **level of significance** then we say **"we have insufficient evidence to reject H_0"**.