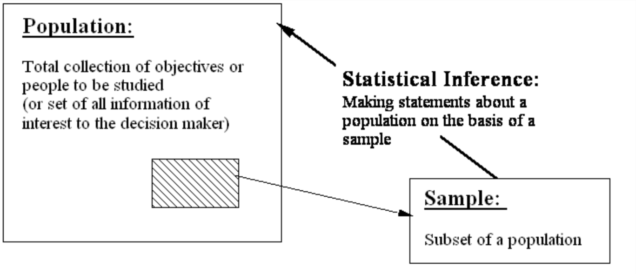
Introduction to Econometrics

Fall 2024

Department of Economics
San Diego State University

Population and Sample



Population Parameter		Sample Parameter
μ_X	mean	$\bar{X}$
σ_X^2	variance	s_X^2
N	size	n

- A sample is a subset of people, items, or events from a larger population that you collect and analyze to make inferences.
- If you're interested in the proportion of red MM's, you likely wouldn't count them all!

Roadmap

- How to understand the data $\implies$ random sample
- What to learn from data $\implies$ parameter
- How to learn from data $\implies$ estimator
- What is the best way to use the data $\implies$ properties of the estimators
- Chapters 1-3 in SW

Random sampling

- To be useful, a sample must be *representative* of the population.
- Simple Random Sample (SRS)
  - A sample is a simple random sample (SRS) if each individual in the population is equally likely to be included in the sample
- Example: class (heights) as a sample of current SDSU undergrads
  - Is this a SRS?
  - If not, how would you construct one?
- An iid (independent and identically distributed) sample

  A sample is an iid random sample if it is a SRS and

  1. each of the n observations in the sample, $X_1, X_2, ..., X_n$, are independent (e.g. no relatives), and
  2. each observation is drawn from the same overall population distribution.

Point estimates

- Samples provide information about population parameters.
- A **point estimate** is a single number that is used to estimate an unknown population parameter.
- The sample mean $\bar{X}$ can be used to estimate the population mean μ_X.
- The sample variance s^2 can be used to estimate the population variance σ_X^2.
- A point *estimate* is a single number.
- An *estimator* is a function of a sample of data that produces that number.

Point estimators

- Formulas for commonly used point estimators:
  - Sample Mean: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}[X_1 + X_2 + ... + X_n]$;
  - Sample Variance: $s_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$.
    - Why $(n-1)$? It comes from using $\bar{X}$ to estimate μ, which introduces a downward bias in $(X_i - \bar{X})^2$ which we correct by multiplying by $\frac{n}{n-1}$(which is > 1)
  - Sample Covariance: $s_{XY} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \overline{Y})$
  - Sample Correlation: $r_{XY} = \frac{s_{XY}}{s_X s_Y}$

Sample mean as a random variable

- The act of random sampling makes the sample mean *itself* a random variable.
- Since the sample mean can be viewed as a random variable (before we actually draw the sample, we don't know what it is going to be), it has a distribution which we call the sampling distribution.
- Thus, the distribution of the sample mean itself has a mean and a standard deviation.
- Let's see what they are...

Distribution of the sample mean

- The expectation of the sample mean for an iid sample of size n from a population with mean μ_X is

$$\begin{aligned} E(\bar{X}) &= E(\tfrac{1}{n}[X_1 + X_2 + ... + X_n]) \\ &= \tfrac{1}{n}[E(X_1) + E(X_2) + ... + E(X_n)] \\ &= \tfrac{1}{n}[\mu_X + \mu_X + ... + \mu_X] \\ &= \tfrac{1}{n}[n\mu_X] \\ &= \mu_X. \end{aligned}$$

- The variance of the sample mean for an iid sample of size n from a population with variance σ_X^2 is

$$\begin{aligned} Var(\bar{X}) &= Var(\tfrac{1}{n}[X_1 + X_2 + ... + X_n]) \\ &= \tfrac{1}{n^2}Var(X_1 + X_2 + ... + X_n) \\ &= \tfrac{1}{n^2}[Var(X_1) + Var(X_2) + ... + Var(X_n)] \\ &= \tfrac{1}{n^2}[\sigma_X^2 + \sigma_X^2 + ... + \sigma_X^2] \quad (iid) \\ &= \tfrac{1}{n^2}[n\sigma_X^2] \\ &= \tfrac{\sigma_X^2}{n}. \end{aligned}$$

Distribution of the sample mean

- The standard deviation of the sample mean is called the **standard error**. The standard error for a sample of size n is

$$SE\left[\bar{X}\right] = \frac{\sigma_X}{\sqrt{n}}$$

or

$$\sigma_{\overline{X}} = \frac{\sigma_X}{\sqrt{n}}$$

Distribution of the sample mean

- Finally, when $X_1, ..., X_n$ are iid draws from the normal distribution $N(\mu_X, \sigma_X^2)$ then

$$\bar{X} \sim N(\mu_X, \frac{\sigma_X^2}{n})$$

- Why would we know that the X_i's were drawn from a normal distribution?
  - We probably wouldn't!
- So what should we do when we don't know what distribution the X_i's are drawn from?

Approximation

- Much of what we do here is based on the "approximate approach". What does this mean?
- Sometimes we will know the exact distribution of $\overline{X}$ such as when the X_i's are iid draws from a normal distribution
- However in many cases we will not know what distribution the X_i's are drawn from, in which case the distribution of $\overline{X}$ will not be known
    - Even if we do know the exact distribution of X, the exact distribution of $\overline{X}$ is likely to be very complicated
- We need to at least *approximate* this distribution in order to build confidence intervals and conduct tests

Approximation

- The approximate approach uses approximations to the sampling distribution that rely on the sample size being large
- This approach relies on two important results from mathematical statistics: the Law of Large Numbers and the Central Limit Theorem
  - LLN: when n is large, $\overline{X}$ will be close to μ_X with high probability (also known as consistency)
  - CLT: when n is large, $\frac{\overline{X}-\mu_X}{\sigma_{\overline{X}}}$ is approximately standard normal (for any distribution of X)

plims, Consistency, and the Law of Large Numbers

- Loosely speaking, we say that a sequence of random variables W_n **converges in probability** to a constant θ if W_n becomes closer and closer to θ as $n \to \infty$.
  - This is written as $W_n \xrightarrow{p} \theta$ or $plim\ (W_n) = \theta$
  - Formally, W_n converges in probability to θ if for every $\varepsilon > 0$

$$P\left(|W_n - \theta| > \varepsilon\right) \to 0 \text{ as } n \to \infty$$

- When convergence in probability is applied to an estimator, we call it consistency.

Consistency

- For example, W_n might be an estimator of a parameter θ based on a sample $X_1, X_2, .., X_n$ of size n.
  - e.g. $\overline{X}_n = \frac{1}{n}[X_1 + X_2 + ... + X_n]$ is a consistent estimator of μ_X.[1]
- We say that W_n is a *consistent* estimator of θ if W_n converges in probability to θ when $n \to \infty$.
- In words, consistency means that as n grows, the distribution of W_n becomes more & more concentrated around θ.

[1] Note: $\overline{X}_n = \overline{X}$ from before, I just want to keep track of n.

The Law of Large Numbers (LLN)

- The **Law of Large Numbers** tells us that $\overline{X}_n$ is a consistent estimator of μ_X. Formally,
  - Let $X_1, X_2, .., X_n$ be iid random variables with mean μ_X and $var\left(X_i\right) = \sigma_X^2 < \infty$. Then

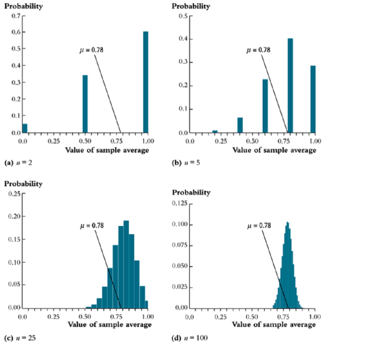
$$plim \ \left(\overline{X}_n\right) = \mu_X \text{ or } \overline{X}_n \xrightarrow{p} \mu_X$$

- When n is large, $\overline{X}_n$ will be close to μ_X with high probability.
- The LLN is sometimes called the law of *averages* because, when you average a large number of random variables with the same mean, the large and small values balance out.
- The proof follows easily from Chebychev's inequality.[2]
- The LLN is illustrated in the example on the next slide.

[2] $\Pr\left(\left|\bar{X} - \mu_X\right| \geq \varepsilon\right) \leq var\left(\bar{X}\right)/\varepsilon^2$ for all $\varepsilon > 0$.

Think of an “unfair” coin flip where $p = P(X_i = 1) = .78$ is the probability of heads.

FIGURE 2.6 Sampling Distribution of the Sample Average of n Bernoulli Random Variables



The distributions are the sampling distributions of $\overline{Y}$, the sample average of n independent Bernoulli random variables with $p = \Pr(Y_i = 1) = 0.78$ (the probability of a fast commute is 78%). The variance of the sampling distribution of $\overline{Y}$ decreases as n gets larger, so the sampling distribution becomes more tightly concentrated around its mean $\mu = 0.78$ as the sample size n increases.

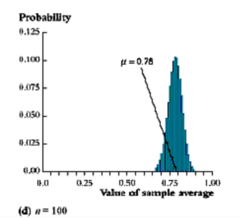
”tails twice” $\overline{X} = 0 \rightarrow P(0)P(0) = (.22)\,(.22) \approx .05$

”heads and tails” $\overline{X} = .5 \rightarrow 2P(0)P(1) = 2\,(.22)\,(.78) \approx .34$

”heads twice” $\overline{X} = 1 \rightarrow P(1)P(1) = (.78)\,(.78) \approx .61$

Central Limit Theorem

- Notice that in the LLN example, when $n = 100$ the sampling distribution looks a lot like a normal distribution.
- This is not just simply a coincidence, but an illustration of an important result called the **Central Limit Theorem** (CLT).
- The CLT states that the average from a random sample for any population (with finite variance), when standardized, has an asymptotic standard normal distribution, meaning that it becomes well approximated by a standard normal.



Central Limit Theorem

- Formally, let $X_1, X_2, .., X_n$ be a random sample with $E\left(X_i\right) = \mu_X$ and $Var\left(X_i\right) = \sigma_X^2 < \infty$. Then as $n \to \infty$

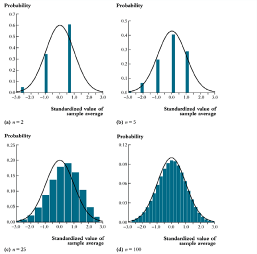
$$\frac{\overline{X} - \mu_X}{\frac{\sigma_X}{\sqrt{n}}} \xrightarrow{d} N\left(0, 1\right) \quad \text{or} \quad \sqrt{n}\left(\overline{X} - \mu_X\right) \xrightarrow{d} N\left(0, \sigma_X^2\right)$$

- This is also sometimes written as

$$\overline{X} \overset{a}{\sim} N\left(\mu, \frac{\sigma_X^2}{n}\right)$$

- When is n large enough to use this approximation?
  - If $X \sim N$: any n
  - If the distribution of X is very non-normal: above 100
- The CLT is illustrated in the following slide.

FIGURE 2.7 Distribution of the Standardized Sample Average of n Bernoulli Random Variables with $p = .78$



The sampling distribution of $\bar{Y}$ in Figure 2.6 is plotted here after standardizing $\bar{Y}$. This centers the distributions in Figure 2.6 and magnifies the scale on the horizontal axis by a factor of $\sqrt{n}$. When the sample size is large, the sampling distributions are increasingly well approximated by the normal distribution (the solid line), as predicted by the central limit theorem.

Recall that for a Bernoulli, $Var\left(X\right) = \sigma_X^2 = p\left(1-p\right)$

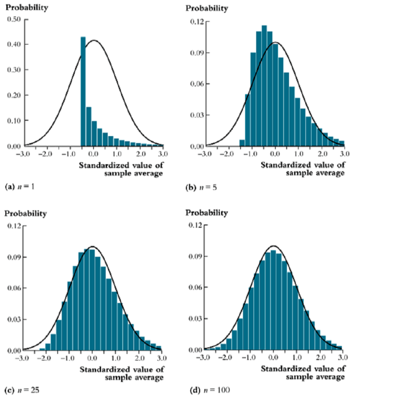
when $\overline{X} = 0: \frac{\overline{X}-\mu}{\sigma_{\overline{X}}} = \frac{0-.78}{\sqrt{\frac{(.78)(.22)}{2}}} \approx -2.66$

when $\overline{X} = .5: \frac{\overline{X}-\mu}{\sigma_{\overline{X}}} = \frac{.5-.78}{\sqrt{\frac{(.78)(.22)}{2}}} \approx -.95$

when $\overline{X} = 1: \frac{\overline{X}-\mu}{\sigma_{\overline{X}}} = \frac{1-.78}{\sqrt{\frac{(.78)(.22)}{2}}} \approx .75$

It even holds for very skewed distributions!

FIGURE 2.8 Distribution of the Standardized Sample Average of *n* Draws from a Skewed Distribution



The figures show the sampling distribution of the standardized sample average of n draws from the skewed (asymmetric) population distribution shown in Figure 2.8a. When n is small ($n = 5$), the sampling distribution, like the population distribution, is skewed. But when n is large ($n = 100$), the sampling distribution is well approximated by a standard normal distribution (solid line), as predicted by the central limit theorem.

Properties of Estimators

- In practice, there are many possible estimators
  - $\overline{X}, x_1, \frac{1}{3}(x_1 + x_5 + x_{22}), \frac{1}{n}\sum_{i=1}^{n}(X_i + 3), \ldots$
- So how should we choose among them?
- We'd like to choose the one that gets as close as possible (at least on average) to the true value.
- More precisely, we would like the sampling distribution to be as tightly centered around the true value as possible.
- In particular, three very desirable characteristics are:
  - Unbiasedness
  - Consistency
  - Efficiency

Unbiasedness

- An estimator W of θ is unbiased if $E(W) = \theta$ ("true on average")
- This doesn't mean that an estimate for a particular sample will equal θ, or even be very close to it.
- However, if we took an average of estimates over all possible samples indefinitely, we would obtain θ.
- For an estimator that is not unbiased, we define its bias as

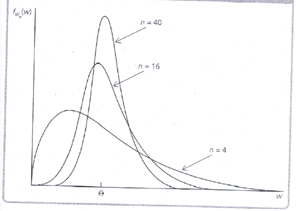
$$Bias(W) \equiv E[W - \theta] = E(W) - \theta$$

- $\bar{X}$ is an unbiased estimator of μ_X,
- s_X^2 is an unbiased estimator of σ_X^2.

Consistency

- An estimator W_n of θ is consistent if $plim\,(W_n) = \theta$ ("true eventually").
- Look again at Figure 2.6 or at the figure below.
- $\bar{X}$ is a consistent estimator of μ_X.
- s_X^2 is a consistent estimator of σ_X^2.[3]

Figure C.3
The sampling distributions of a consistent estimator for three sample sizes.



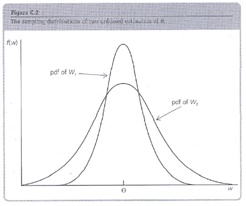
[3]This is proven (using the Law of Large Numbers) in appendix 3.3 of S&W.

Consistency

- Econometricians typically care more about consistency than bias
- A biased but consistent estimator may not be true on average, but it will be closer to the truth as n gets bigger
- An unbiased but inconsistent estimator can still be far from the truth for very large n

Efficiency

- Suppose you have two unbiased estimators (and a fixed sample size). How should you choose between them?
- One option is to choose the one with the tightest sampling distribution.
- Comparing estimators on the basis of variance is called efficiency:
- If W_1 and W_2 are two unbiased estimators of θ, then W_1 is more efficient than W_2 if $Var\,(W_1) \leq Var\,(W_2)$ for all θ.
- Example: $\frac{1}{3}(x_1 + x_2 + x_3)$ and $\frac{1}{4}x_1 + \frac{1}{2}x_2 + \frac{1}{4}x_3$



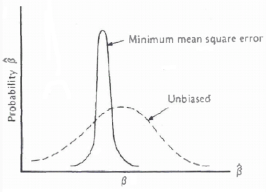
Other Desirable Properties

- Mean Squared Error (MSE)
  - If we don't restrict our attention to unbiased estimators, then comparing variances is meaningless.
    - Example: estimating μ with 0
  - Also, we might prefer a slightly biased but efficient estimator to an unbiased but inefficient one.
- One way to compare estimators that are not necessarily unbiased is to compute the MSE

$$MSE\,(W) = E\left[(W-\theta)^2\right]$$

Mean Square Error

- The MSE measures how far, on average, the estimator is away from θ
  - It can be shown that $MSE(W) = Var(W) + [Bias(W)]^2$
    - $MSE\,(W)$ depends on both the variance and bias
  - Intuitively we want the MSE to be small. Note that if $Bias = 0$ then we just want to minimize the variance of the estimator



Confidence Intervals and Hypothesis Testing

- For a given parameter, we might be interested in:
    1. Reporting a range of values which you believe will contain the true value of the parameter with high probability.
    2. Determining whether the data support or reject some prior belief about the parameter (is it greater than 0? equal to 19?).
- Since we can approximate the distribution of the estimator, we can do both!

Confidence Intervals

- A point estimate by itself does not provide enough information for testing economic theories or informing policy discussions.
- A point estimate is a researcher's best guess at the population value, but it doesn't tell us how close the estimate is likely to be to the population parameter.
- Therefore, an estimate is much more useful if it is accompanied by an estimate of the error that might be involved, that is:

$$\mu_X = \bar{X} \pm error$$

or

$$\bar{X} - error < \mu_X < \bar{X} + error.$$

Interval estimate

- An **interval estimate** describes a range of values within which a population parameter is likely to lie

$$\left[\bar{X} - error, \bar{X} + error\right]$$

- Associated with the interval is a measure of the confidence we have that the interval does indeed contain the parameter of interest.

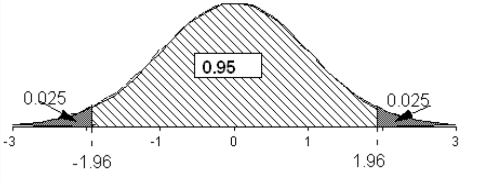
$$\Pr(\bar{X} - error < \mu_X < \bar{X} + error) = 0.95$$

- Why $.95$?

Constructing a Confidence Interval

- Let's start with a closely related exercise.
- Find an interval that will contain a random draw from a standard normal distribution 95% of the time
  - From the Normal tables we know

$$\Pr(-1.96 < Z < 1.96) = 0.95$$



- There is a 95% chance that the random variable Z will fall in the interval $(-1.96, 1.96)$.
- Or, in other words, our best guess for the value of that draw is 0 ± 1.96

Constructing a Confidence Interval

- Now let's construct a confidence interval for a population parameter
- How about a 95% CI for the mean (μ_X)?
  - Our best guess (point estimate) is $\bar{X}$.
  - We already know (from the CLT) that $\frac{\bar{X}-\mu_X}{\frac{\sigma_X}{\sqrt{n}}}$ is (approximately) $N(0,1)$.
  - Therefore $\Pr\left(-1.96 < \frac{\bar{X}-\mu_X}{\frac{\sigma_X}{\sqrt{n}}} < 1.96\right) = 0.95$ as above.

Constructing a Confidence Interval

$$\Pr\left(-1.96 < \frac{\bar{X} - \mu_X}{\frac{\sigma_X}{\sqrt{n}}} < 1.96\right) = 0.95$$

- Rearranging,

$$\Pr\left(\bar{X} - 1.96 \cdot \frac{\sigma_X}{\sqrt{n}} < \mu_X < \bar{X} + 1.96 \cdot \frac{\sigma_X}{\sqrt{n}}\right) = 0.95.$$

- Thus, a 95% confidence interval for μ_X is

$$\mu_X = \bar{X} \pm 1.96 \cdot \frac{\sigma_X}{\sqrt{n}} = \left[\bar{X} - 1.96 \cdot \frac{\sigma_X}{\sqrt{n}}, \bar{X} + 1.96 \cdot \frac{\sigma_X}{\sqrt{n}}\right]$$

- Notice that the interval gets tighter as $n \longrightarrow \infty$

Interpretation

$$\mu_X = \bar{X} \pm 1.96 \cdot \frac{\sigma_X}{\sqrt{n}} = \left[\bar{X} - 1.96 \cdot \frac{\sigma_X}{\sqrt{n}}, \bar{X} + 1.96 \cdot \frac{\sigma_X}{\sqrt{n}}\right]$$

- Which statement about the interval above is correct?
  1. "There is a 95% chance that the population mean μ_X is contained in the interval."
  2. "There is a 95% chance that the interval will contain the population mean μ_X."
- The answer is 2:
  - Population parameters are constants, confidence intervals are random.

Example - Call Duration

- Suppose you are the consumer helpline manager at Microsoft and you are interested in finding out how long your customer service people spend on the phone
- You observe 140 calls, and the sample mean is $\bar{X} = 3.91$.
- Suppose that we *know* that the population standard deviation, σ_X, is 7.8.
  - Why would we? Let's just pretend for now...
- Then the sample standard deviation is $\frac{\sigma_X}{\sqrt{n}} = \frac{7.8}{\sqrt{140}} = 0.66$.
- Calculate a 95% confidence interval for call duration (suppose you want to be 95% sure your interval contains the true mean)

$$\mu_X = \overline{X} \pm 1.96 \cdot \frac{\sigma_X}{\sqrt{n}} = 3.91 \pm 1.96 \cdot .66 = 3.91 \pm 1.29 = (2.62, 5.20)$$

General Format for Confidence Intervals

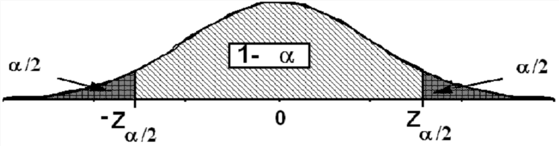
- Of course, we might not always want to construct a 95% CI, what if we want a 90% or 99% CI?

General Format for Confidence Intervals

- Of course, we might not always want to construct a 95% CI, what if we want a 90% or 99% CI?
- We need to replace 1.96 with the appropriate number.

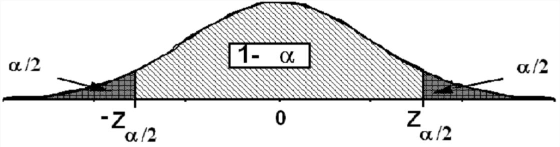
General Format for Confidence Intervals

- Of course, we might not always want to construct a 95% CI, what if we want a 90% or 99% CI?
- We need to replace 1.96 with the appropriate number.
- Let α be the probability that the CI doesn't include the true value (and $1-\alpha$ the probability that it does)



General Format for Confidence Intervals

- Of course, we might not always want to construct a 95% CI, what if we want a 90% or 99% CI?
- We need to replace 1.96 with the appropriate number.
- Let α be the probability that the CI doesn't include the true value (and $1-\alpha$ the probability that it does)



- So we want to find the boundary points $\left(-z_{\alpha/2}, z_{\alpha/2}\right)$ such that the probability that the CI does contain the true value equals $1-\alpha$.

- Since $\frac{\bar{X}-\mu_X}{\frac{\sigma_X}{\sqrt{n}}} \sim N(0,1)$, which is symmetric, this is the same as finding $z_{\alpha/2}$ such that

$$\Pr(Z \leq -z_{\alpha/2}) = \alpha/2.$$

- Since $\frac{\bar{X}-\mu_X}{\frac{\sigma_X}{\sqrt{n}}} \sim N(0,1)$, which is symmetric, this is the same as finding $z_{\alpha/2}$ such that

$$\Pr(Z \leq -z_{\alpha/2}) = \alpha/2.$$

- For example, if $\alpha = .05$ then $z_{\alpha/2} = 1.96$ (since $\Phi(-1.96) = .025$), but if $\alpha = .1$ then $z_{\alpha/2} = 1.645$ (since $\Phi(-1.645) = .05$).

	Second Decimal Value of z									
z	0	1	2	3	4	5	6	7	8	9
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611

General Format

- Of course, this is just restating the fact the probability that Z falls between $-z_{\alpha/2}$ and $z_{\alpha/2}$ is

$$\Pr(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha$$

General Format

- Of course, this is just restating the fact the probability that Z falls between $-z_{\alpha/2}$ and $z_{\alpha/2}$ is

$$\Pr(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha$$

- Since $\frac{\bar{X}-\mu_X}{\frac{\sigma_X}{\sqrt{n}}} = Z$

$$\Pr(-z_{\alpha/2} < \frac{\bar{X}-\mu_X}{\frac{\sigma_X}{\sqrt{n}}} < z_{\alpha/2}) = 1 - \alpha$$

General Format

- Of course, this is just restating the fact the probability that Z falls between $-z_{\alpha/2}$ and $z_{\alpha/2}$ is

$$\Pr(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha$$

- Since $\frac{\bar{X}-\mu_X}{\frac{\sigma_X}{\sqrt{n}}} = Z$

$$\Pr(-z_{\alpha/2} < \frac{\bar{X} - \mu_X}{\frac{\sigma_X}{\sqrt{n}}} < z_{\alpha/2}) = 1 - \alpha$$

- By rearranging the terms, we get

$$\Pr(\bar{X} - z_{\alpha/2} \cdot \frac{\sigma_X}{\sqrt{n}} < \mu_X < \bar{X} + z_{\alpha/2} \cdot \frac{\sigma_X}{\sqrt{n}}) = 1 - \alpha.$$

General Format

- Thus, a $100(1-\alpha)\%$ confidence interval for μ_X is

$$\mu_X = \bar{X} \pm z_{\alpha/2} \cdot \frac{\sigma_X}{\sqrt{n}} = \left[\bar{X} - z_{\alpha/2} \cdot \frac{\sigma_X}{\sqrt{n}}, \bar{X} + z_{\alpha/2} \cdot \frac{\sigma_X}{\sqrt{n}}\right]$$

General Format

- Thus, a $100(1-\alpha)\%$ confidence interval for μ_X is

$$\mu_X = \bar{X} \pm z_{\alpha/2} \cdot \frac{\sigma_X}{\sqrt{n}} = \left[\bar{X} - z_{\alpha/2} \cdot \frac{\sigma_X}{\sqrt{n}}, \bar{X} + z_{\alpha/2} \cdot \frac{\sigma_X}{\sqrt{n}}\right]$$

- Note that by picking α you are choosing the value of $z_{\alpha/2}$ in this interval, which changes the size of the interval (as α decreases the interval gets larger).

General Format

- Thus, a $100(1-\alpha)\%$ confidence interval for μ_X is

$$\mu_X = \bar{X} \pm z_{\alpha/2} \cdot \frac{\sigma_X}{\sqrt{n}} = \left[\bar{X} - z_{\alpha/2} \cdot \frac{\sigma_X}{\sqrt{n}}, \bar{X} + z_{\alpha/2} \cdot \frac{\sigma_X}{\sqrt{n}}\right]$$

- Note that by picking α you are choosing the value of $z_{\alpha/2}$ in this interval, which changes the size of the interval (as α decreases the interval gets larger).
- This means that you are *choosing* how often (on average) you are willing to be wrong!

CI for the mean when the SD is unknown

- So far we have been assuming that we know σ_X.
- Of course, when the mean of the population (μ_X) is unknown and must be estimated, we probably don't know σ_X either.

CI for the mean when the SD is unknown

- So far we have been assuming that we know σ_X.
- Of course, when the mean of the population (μ_X) is unknown and must be estimated, we probably don't know σ_X either.
- Instead, the population standard deviation must be estimated from the sample standard deviation s_X where

$$s_X^2 = \frac{1}{n-1} \sum \left(X_i - \bar{X} \right)^2$$

CI for a the mean when the SD is unknown

- Therefore, instead of $Z = \frac{\bar{X} - \mu_X}{\frac{\sigma_X}{\sqrt{n}}}$, we use the *t*-statistic

$$t = \frac{\bar{X} - \mu_X}{\frac{s_X}{\sqrt{n}}}$$

where $\frac{s_X}{\sqrt{n}} \equiv SE\left(\bar{X}\right)$

CI for a the mean when the SD is unknown

- Therefore, instead of $Z = \frac{\bar{X}-\mu_X}{\frac{\sigma_X}{\sqrt{n}}}$, we use the t-statistic

$$t = \frac{\bar{X} - \mu_X}{\frac{s_X}{\sqrt{n}}}$$

  where $\frac{s_X}{\sqrt{n}} \equiv SE\left(\bar{X}\right)$

- Since t is approximately distributed $N\left(0,1\right)$ for large n, a $100(1-\alpha)\%$ confidence interval for μ is simply

$$\mu_X = \bar{X} \pm z_{\alpha/2} \cdot \frac{s_X}{\sqrt{n}} = \left[\bar{X} - z_{\alpha/2} \cdot \frac{s_X}{\sqrt{n}}, \bar{X} + z_{\alpha/2} \cdot \frac{s_X}{\sqrt{n}}\right]$$

Hypothesis Testing: The Basic Idea

- We've seen now how to construct and interpret confidence intervals and how to use them to quantify the accuracy of our estimate.
- How else can they be used?

Hypothesis Testing: The Basic Idea

- We've seen now how to construct and interpret confidence intervals and how to use them to quantify the accuracy of our estimate.
- How else can they be used?
- Sometimes a question we are interested in has a definite yes or no answer.
  - Does a job training program increase worker productivity?
  - Are women discriminated against in hiring?
  - Does having more police officers reduce the crime rate?

Hypothesis Testing: The Basic Idea

- We've seen now how to construct and interpret confidence intervals and how to use them to quantify the accuracy of our estimate.
- How else can they be used?
- Sometimes a question we are interested in has a definite yes or no answer.
  - Does a job training program increase worker productivity?
  - Are women discriminated against in hiring?
  - Does having more police officers reduce the crime rate?
- Answering these questions, *using a sample of data*, is called hypothesis testing.
- CIs are one of several ways to test hypotheses.

Hypothesis Testing: Some terminology

- *Hypotheses* are assumptions about a population parameter.
  - $\mu_X = 21,600$

Hypothesis Testing: Some terminology

- *Hypotheses* are assumptions about a population parameter.
  - $\mu_X = 21,600$
- *Hypothesis testing* is a process enabling the correctness of hypotheses to be examined on the basis of *sample information.*
  - $\bar{X} = 25,000$ is simply *too large* to accept this hypothesis

The General Process

- Hypotheses come in pairs that, together, describe all the relevant possibilities.
  - Null hypothesis (H_0) – e.g. $H_0 : \mu_X = 21,600$
  - Alternative hypothesis[4] (H_A) – e.g. $H_A : \mu_X \neq 21,600$

[4] The alternative hypothesis is sometimes also written as H_1.

The General Process

- Hypotheses come in pairs that, together, describe all the relevant possibilities.
  - Null hypothesis (H_0) – e.g. $H_0 : \mu_X = 21,600$
  - Alternative hypothesis[4] (H_A) – e.g. $H_A : \mu_X \neq 21,600$
- H_0 is like the defendant in a trial: the null hypothesis is assumed true until the data strongly suggests otherwise.
- The hypothesis testing procedure focuses on whether H_0 is consistent with the data.

[4] The alternative hypothesis is sometimes also written as H_1.

The General Process

- Hypotheses come in pairs that, together, describe all the relevant possibilities.
  - Null hypothesis (H_0) – e.g. $H_0 : \mu_X = 21,600$
  - Alternative hypothesis[4] (H_A) – e.g. $H_A : \mu_X \neq 21,600$
- H_0 is like the defendant in a trial: the null hypothesis is assumed true until the data strongly suggests otherwise.
- The hypothesis testing procedure focuses on whether H_0 is consistent with the data.
- When sample information (e.g. $\bar{X}$) suggests that H_0 is incorrect, H_0 is said to be <u>rejected</u>.

[4] The alternative hypothesis is sometimes also written as H_1.

Interpretation

- Interpretation of rejecting H_0?
  - H_A is more consistent with the data.

Interpretation

- Interpretation of rejecting H_0?
  - H_A is more consistent with the data.
- What if you do not reject H_0?
- Should you say that you Accept H_0 or that you Cannot Reject H_0?

Interpretation

- Interpretation of rejecting H_0?
  - H_A is more consistent with the data.
- What if you do not reject H_0?
- Should you say that you Accept H_0 or that you Cannot Reject H_0?
  - Rule: you always say you Cannot Reject H_0
  - Why? There will always be lots of null hypotheses that you can't reject! (Example: $H_0 : \mu_X = 21,601$)

Determining the Null and Alternative

1. Start by using a null hypothesis with an "=" sign. Use $H_0 : \mu_X =$ [the pivotal value].

2. What do you want to try to prove? Use that as the alternative hypothesis. (This is the standard convention)

One-sided versus Two-sided Hypotheses

- The rejection rule for a null hypothesis will depend on the nature of the alternative.

One-sided versus Two-sided Hypotheses

- The rejection rule for a null hypothesis will depend on the nature of the alternative.
- Alternative hypotheses can be either one- or two-sided

$$
\begin{aligned}
H_A &: \mu_X > \mu_0 \\
H_A &: \mu_X < \mu_0
\end{aligned}
$$

or

$$H_A : \mu_X \neq \mu_0$$

One-sided versus Two-sided Hypotheses: Examples

- Two-sided
  - Example: Is a new TV manufacturing process <u>different</u> from the old? i.e. it might be better or worse.

$$H_0 : \mu_X = 1200,\ H_A : \mu_X \neq 1200$$

  - H_0 always involves an "$=$" sign. In a two-sided test, H_A involves a "$\neq$" sign.
- Intuition: Either a big **or** small $\bar{X}$ is evidence against the null.

One-sided versus Two-sided Hypotheses: Examples

- One-sided
  - Example: You know the new manufacturing process is not worse than the old. But is it better?

$$H_0 : \mu_X = 1200, \quad H_A : \mu_X > 1200$$

  - H_0 always involves an "$=$" sign. In a one-sided hypothesis test, H_A will always have a "$<$" or "$>$" sign.
- Intuition: Only a big $\bar{X}$ is evidence against the null.

One-sided versus Two-sided Hypotheses

- In practice, one-sided alternative hypotheses should be used when there is a clear reason for μ being on a certain side of the null value under the alternative.
  - For example, smaller classes seem likely to improve student's test scores since they improve the learning environment.

One-sided versus Two-sided Hypotheses

- In practice, one-sided alternative hypotheses should be used when there is a clear reason for μ being on a certain side of the null value under the alternative.
  - For example, smaller classes seem likely to improve student's test scores since they improve the learning environment.
- However, in practice, most alternatives are two-sided because it's typically difficult to be certain about the direction of an effect.
  - For example, a drug which is meant to improve one's health could have harmful side effects.

Hypothesis Testing: 3 Methods

- So how do we determine which hypothesis is more consistent with the data?
    - Intuition: we determine whether $\bar{X}$ & μ_0 are "too far apart".
- There are three **completely equivalent** methods for doing this:
    - Confidence Intervals
    - Acceptance Regions
    - p-values

Confidence intervals

- Consider a two-sided null hypothesis.
- Since under the null $t = \frac{\bar{X} - \mu_X}{\frac{s_X}{\sqrt{n}}} \sim N(0,1)$,

$$\Pr\left(\bar{X} - z_{\alpha/2}\frac{s_X}{\sqrt{n}} < \mu_X < \bar{X} + z_{\alpha/2}\frac{s_X}{\sqrt{n}}\right) = 1 - \alpha.$$

- Thus a $100(1-\alpha)\%$ confidence interval around $\bar{X}$ is

$$\left(\bar{X} - z_{\alpha/2}\frac{s_X}{\sqrt{n}},\ \bar{X} + z_{\alpha/2}\frac{s_X}{\sqrt{n}}\right)$$

Confidence intervals

- A $100(1-\alpha)\%$ confidence interval around $\bar{X}$ is

$$\left(\bar{X} - z_{\alpha/2}\frac{s_X}{\sqrt{n}},\ \bar{X} + z_{\alpha/2}\frac{s_X}{\sqrt{n}}\right)$$

- This is the range that should include μ_X $100(1-\alpha)\%$ of the time (on average).

Confidence intervals

- A $100(1-\alpha)\%$ confidence interval around $\bar{X}$ is

$$\left(\bar{X} - z_{\alpha/2}\frac{s_X}{\sqrt{n}},\ \bar{X} + z_{\alpha/2}\frac{s_X}{\sqrt{n}}\right)$$

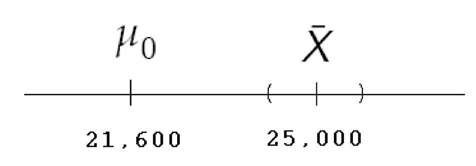
- This is the range that should include μ_X $100(1-\alpha)\%$ of the time (on average).
- Decision rule: reject H_0 if the value of μ_X in the null hypothesis (i.e. μ_0) lies outside the confidence interval; otherwise, do not reject.

Confidence intervals

- A $100(1-\alpha)\%$ confidence interval around $\bar{X}$ is

$$\left(\bar{X} - z_{\alpha/2}\frac{s_X}{\sqrt{n}},\ \bar{X} + z_{\alpha/2}\frac{s_X}{\sqrt{n}}\right)$$

- This is the range that should include μ_X $100(1-\alpha)\%$ of the time (on average).
- Decision rule: reject H_0 if the value of μ_X in the null hypothesis (i.e. μ_0) lies outside the confidence interval; otherwise, do not reject.
- Intuition: the confidence interval is the set of all null hypotheses that you cannot reject at the $\alpha\%$ level.



Acceptance regions

- Consider a two-sided null hypothesis

$$\begin{aligned} H_0 &: \mu_X = \mu_0 \\ H_A &: \mu_X \neq \mu_0 \end{aligned}$$

- Since under the null, $t = \frac{\bar{X} - \mu_0}{\frac{s_X}{\sqrt{n}}} \sim N(0,1)$,

$$\Pr\left(\bar{X} - z_{\alpha/2}\frac{s_X}{\sqrt{n}} < \mu_0 < \bar{X} + z_{\alpha/2}\frac{s_X}{\sqrt{n}}\right) = 1 - \alpha.$$

Acceptance regions

- Consider a two-sided null hypothesis

$$\begin{aligned} H_0 &: \mu_X = \mu_0 \\ H_A &: \mu_X \neq \mu_0 \end{aligned}$$

- Since under the null, $t = \frac{\bar{X} - \mu_0}{\frac{s_X}{\sqrt{n}}} \sim N(0, 1)$,

$$\Pr\left(\bar{X} - z_{\alpha/2}\frac{s_X}{\sqrt{n}} < \mu_0 < \bar{X} + z_{\alpha/2}\frac{s_X}{\sqrt{n}}\right) = 1 - \alpha.$$

- This can easily be rearranged as

$$\Pr\left(\mu_0 - z_{\alpha/2}\frac{s_X}{\sqrt{n}} < \bar{X} < \mu_0 + z_{\alpha/2}\frac{s_X}{\sqrt{n}}\right) = 1 - \alpha.$$

Acceptance regions

- Thus a $100(1-\alpha)\%$ acceptance region around μ_0 is

$$\left(\mu_0 - z_{\alpha/2}\frac{s_X}{\sqrt{n}},\ \mu_0 + z_{\alpha/2}\frac{s_X}{\sqrt{n}}\right)$$

Acceptance regions

- Thus a $100(1-\alpha)\%$ acceptance region around μ_0 is

$$\left(\mu_0 - z_{\alpha/2}\frac{s_X}{\sqrt{n}},\ \mu_0 + z_{\alpha/2}\frac{s_X}{\sqrt{n}}\right)$$

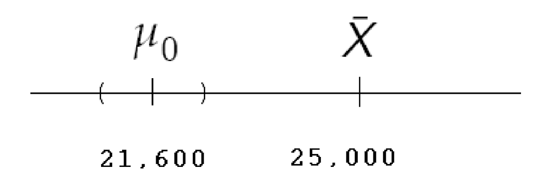
- You can think of this as the range that $\bar{X}$ should lie in if H_0 is true.

Acceptance regions

- Thus a $100(1-\alpha)\%$ acceptance region around μ_0 is

$$\left(\mu_0 - z_{\alpha/2}\frac{s_X}{\sqrt{n}},\ \mu_0 + z_{\alpha/2}\frac{s_X}{\sqrt{n}}\right)$$

- You can think of this as the range that $\bar{X}$ should lie in if H_0 is true.
- Decision Rule: reject H_0 if $\bar{X}$ lies outside the acceptance region; otherwise, do not reject.



Example: TVs, Two-Sided Alternative

Suppose $\overline{X} = 1265$

1. Establish null and alternative hypotheses

   $H_0: \quad \mu_X = 1200 \qquad\qquad H_A: \quad \mu_X \neq 1200$

2. Choose a significance level

   $\alpha = .05$

3. Calculate a 95% acceptance region $(s_X = 300,\ n = 100)$

   $AR = \left(\mu_0 - z_{\alpha/2} \cdot \frac{s_X}{\sqrt{n}}, \mu_0 + z_{\alpha/2} \cdot \frac{s_X}{\sqrt{n}}\right) =$

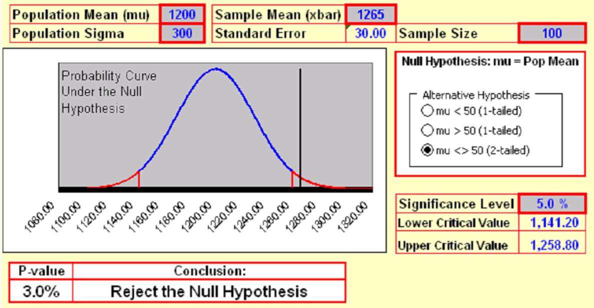
   $(1200 - 1.96 \cdot 30, 1200 + 1.96 \cdot 30) = (1141.2, 1258.8)$

4. Reach a conclusion

   $\overline{X} = 1265$ lies outside AR so you should reject H_0.

TV Example

$$AR = \left(\mu_0 - z_{\alpha/2} \cdot \frac{s_X}{\sqrt{n}}, \mu_0 + z_{\alpha/2} \cdot \frac{s_X}{\sqrt{n}}\right) = (1141.2, 1258.8)$$



Acceptance Regions versus Confidence Intervals

- So our acceptance region was

$$AR = \left(\mu_0 - z_{\alpha/2} \cdot \frac{s_X}{\sqrt{n}}, \mu_0 + z_{\alpha/2} \cdot \frac{s_X}{\sqrt{n}}\right) = (1141.2, 1258.8)$$

which led us to reject H_0.

Acceptance Regions versus Confidence Intervals

- So our acceptance region was

$$AR = \left(\mu_0 - z_{\alpha/2} \cdot \frac{s_X}{\sqrt{n}}, \mu_0 + z_{\alpha/2} \cdot \frac{s_X}{\sqrt{n}}\right) = (1141.2, 1258.8)$$

which led us to reject H_0.

- What if we used a confidence interval instead?

$$CI = \left(\bar{X} - z_{\alpha/2} \cdot \frac{s_X}{\sqrt{n}}, \bar{X} + z_{\alpha/2} \cdot \frac{s_X}{\sqrt{n}}\right) =$$

$$(1265 - 1.96 \cdot 30, 1265 + 1.96 \cdot 30) = (1206.2, 1323.8)$$

so we still reject H_0 (of course!).

Sample Size and Significance Level

- What if $n = 25$? Now $\frac{s_X}{\sqrt{n}} = \frac{300}{5} = 60$

$AR = (1200 \pm 1.96 \cdot 60) = (1200 \pm 117.6) = (1082.4, 1317.6)$

Sample Size and Significance Level

- What if $n = 25$? Now $\frac{s_X}{\sqrt{n}} = \frac{300}{5} = 60$

$AR = (1200 \pm 1.96 \cdot 60) = (1200 \pm 117.6) = (1082.4, 1317.6)$

Now you can't reject H_0! Does this make sense?!?

Sample Size and Significance Level

- What if $n = 25$? Now $\frac{s_X}{\sqrt{n}} = \frac{300}{5} = 60$

$AR = (1200 \pm 1.96 \cdot 60) = (1200 \pm 117.6) = (1082.4, 1317.6)$

Now you can't reject H_0! Does this make sense?!?

- What if $\alpha = .01$? (let's assume $n = 100$ again)

$AR = (1200 \pm 2.58 \cdot 30) = (1200 \pm 77.4) = (1122.6, 1277.4)$

Sample Size and Significance Level

- What if $n = 25$? Now $\frac{s_X}{\sqrt{n}} = \frac{300}{5} = 60$

$AR = (1200 \pm 1.96 \cdot 60) = (1200 \pm 117.6) = (1082.4, 1317.6)$

Now you can't reject H_0! Does this make sense?!?

- What if $\alpha = .01$? (let's assume $n = 100$ again)

$AR = (1200 \pm 2.58 \cdot 30) = (1200 \pm 77.4) = (1122.6, 1277.4)$

Again, you can't reject H_0.

Sample Size and Significance Level

- What if $n = 25$? Now $\frac{s_X}{\sqrt{n}} = \frac{300}{5} = 60$

$AR = (1200 \pm 1.96 \cdot 60) = (1200 \pm 117.6) = (1082.4, 1317.6)$

Now you can't reject H_0! Does this make sense?!?

- What if $\alpha = .01$? (let's assume $n = 100$ again)

$AR = (1200 \pm 2.58 \cdot 30) = (1200 \pm 77.4) = (1122.6, 1277.4)$

Again, you can't reject H_0.

- Both the sample size and the level of significance you choose can matter a lot!

One-Sided Hypotheses

- What about one-sided hypotheses?
- The acceptance regions for one-sided tests are:
  - $H_0 : \mu_X = 1200 \quad H_A : \mu_X > 1200$

$$\left(-\infty,\ \mu_0 + z_\alpha \frac{s_X}{\sqrt{n}}\right)$$

    - Intuition: a big $\bar{X}$ is surprising.

One-Sided Hypotheses

- What about one-sided hypotheses?
- The acceptance regions for one-sided tests are:
  - $H_0 : \mu_X = 1200 \quad H_A : \mu_X > 1200$

$$\left(-\infty,\ \mu_0 + z_\alpha \frac{s_X}{\sqrt{n}}\right)$$

    - Intuition: a big $\bar{X}$ is surprising.
  - $H_0 : \mu_X = 1200 \quad H_A : \mu_X < 1200$

$$\left(\mu_0 - z_\alpha \frac{s_X}{\sqrt{n}},\ \infty\right).$$

    - Intuition: a small $\bar{X}$ is surprising.

One-Sided Hypothesis Testing Example

- Suppose we instead test the one-sided hypothesis:

$$\begin{aligned} H_0 &: \mu_X = 1200 \\ H_A &: \mu_X > 1200 \end{aligned}$$

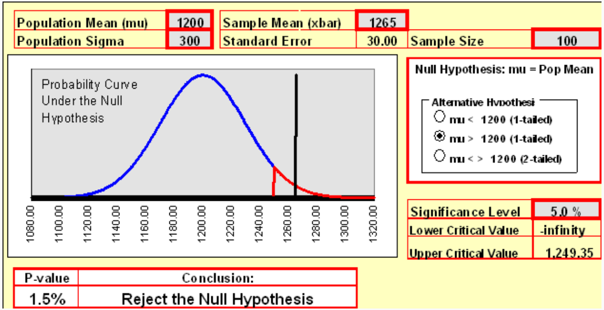
- Again, suppose that $\bar{X} = 1265, s_X = 300$, and $n = 100$
- The acceptance region for $\alpha = .05$ is

$$AR = \left(-\infty, \mu_0 + z_\alpha \frac{s_X}{\sqrt{n}}\right) = \left(-\infty, 1200 + 1.64 \frac{300}{\sqrt{100}}\right) = (-\infty, 1249.35)$$

- Since $\bar{X} = 1265 > 1249.35$, reject H_0 at the 5% level.

One-Sided Example

- The graph below shows the distribution of $\bar{X}$ you expect to see if H_0 is true.
- $\bar{X} = 1265$ is pretty far out in the tail (which is why we reject).



p-Values

- This brings us to the third and most useful procedure for testing hypotheses, the p-value.

p-Values

- This brings us to the third and most useful procedure for testing hypotheses, the p-value.
- Simply put, the p-value is the largest significance level at which we could carry out the test and still fail to reject H_0.

p-Values

- This brings us to the third and most useful procedure for testing hypotheses, the p-value.
- Simply put, the p-value is the largest significance level at which we could carry out the test and still fail to reject H_0.
- What does this mean?

p-Values

- This brings us to the third and most useful procedure for testing hypotheses, the p-value.
- Simply put, the p-value is the largest significance level at which we could carry out the test and still fail to reject H_0.
- What does this mean?
- In the previous example, we clearly could have rejected H_0 at less than the 5% level. But how much less?

p-Values

- For a given sample size, the acceptance region expands as α decreases and shrinks as α increases.

p-Values

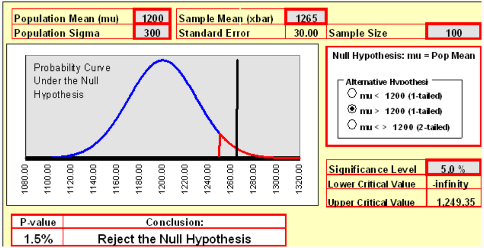
- For a given sample size, the acceptance region expands as α decreases and shrinks as α increases.
- Given the sample average $\bar{X}$, for what α would $\bar{X}$ equal the upper bound of the acceptance region? This is the p-value.

p-Values

- For a given sample size, the acceptance region expands as α decreases and shrinks as α increases.
- Given the sample average $\bar{X}$, for what α would $\bar{X}$ equal the upper bound of the acceptance region? This is the p-value.
- The p-value is the probability of drawing an $\bar{X}$, by random sampling variation, at least as far from the H_0 value as the $\bar{X}$ actually observed (assuming H_0 is correct). Thus, a small value for the p-value is evidence against H_0.

p-Values

- For a given sample size, the acceptance region expands as α decreases and shrinks as α increases.
- Given the sample average $\bar{X}$, for what α would $\bar{X}$ equal the upper bound of the acceptance region? This is the p-value.
- The p-value is the probability of drawing an $\bar{X}$, by random sampling variation, at least as far from the H_0 value as the $\bar{X}$ actually observed (assuming H_0 is correct). Thus, a small value for the p-value is evidence against H_0.
- Decision Rule: reject the null if the p-value is less than α.



p-Values

- For example, for the (one-sided) alternative that the population mean is greater than some number (e.g. $H_A : \mu_X > \mu_0$), the p-value is the value of α such that the upper bound of the acceptance region is equal to $\bar{X}$, i.e. the value of α that satisfies

$$\bar{X} = \mu_0 + z_\alpha \frac{s_X}{\sqrt{n}}.$$

- This is equivalent to finding α such that $z_\alpha = \frac{\bar{X}-\mu_0}{\frac{s_X}{\sqrt{n}}} = t$.
  - The standardized sample average $\left(\frac{\bar{X}-\mu_0}{\frac{s_X}{\sqrt{n}}}\right)$ is called the t-statistic and is denoted t.

p-Values

- So, for an alternative hypothesis that the population mean is large (e.g. $H_A : \mu > 1200$), this is equivalent to finding
  - p-value $= \Pr\left(Z > \frac{\bar{X}-\mu_0}{\frac{s_X}{\sqrt{n}}}\right) = 1 - \Phi\left(t\right).$
- For an alternative hypothesis that the population mean is small (e.g. $H_A : \mu < 1200$), the
  - p-value $= \Pr\left(Z < \frac{\bar{X}-\mu_0}{\frac{s_X}{\sqrt{n}}}\right) = \Phi\left(\frac{\bar{X}-\mu_0}{\frac{s_X}{\sqrt{n}}}\right) = \Phi\left(t\right).$
- For a two-sided alternative hypothesis, the
  - p-value $= 2 \cdot \Pr\left(Z > \left|\frac{\bar{X}-\mu_0}{\frac{s_X}{\sqrt{n}}}\right|\right) = 2 \cdot \Phi\left(-\left|t\right|\right)$

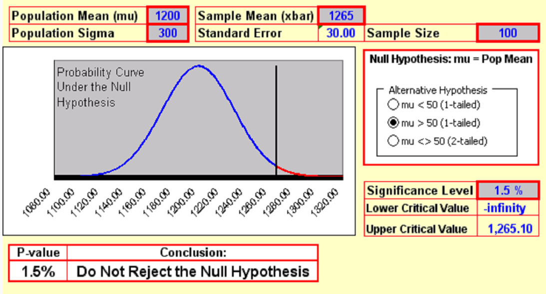
p-Value example

TV's again: $H_0: \quad \mu_X = 1200 \qquad H_A: \quad \mu_X > 1200$

$$p\text{-value} = \Pr\left(Z \geq \frac{\bar{X} - \mu_0}{\frac{s_X}{\sqrt{n}}}\right) = \Pr\left(Z \geq \frac{1265 - 1200}{\frac{300}{\sqrt{100}}}\right)$$

$$= \Pr\left(Z \geq 2.1667\right) = 1 - \Phi\left(2.1667\right)$$

$$= 1 - .985 = 0.015$$



p-Values and Critical Values

- Decision Rule: reject the null if the p-value is less than α.
- An equivalent rule can be formed by comparing t to a critical value.
- **Example**: Two-sided test at a 5% significance level

$$
\begin{aligned}
\text{Reject if } p\text{-value } = 2 \cdot \Phi\left(-\left|t\right|\right) &< 0.05 \\
\Phi\left(-\left|t\right|\right) &< 0.025 \\
-\left|t\right| &< \Phi^{-1}(0.025) = -1.96 \\
\left|t\right| &> 1.96
\end{aligned}
$$

- Here, 1.96 is the critical value for 5% significance level.

p-Values and Critical Values

- **Example**: One-sided test $(H_A : \mu > \mu_0)$ at a 5% significance level

$$\begin{aligned} \text{Reject if } p\text{-value } = 1 - \Phi(t) &< 0.05 \\ -\Phi(t) &< -0.95 \\ t &> \Phi^{-1}(0.95) = 1.645 \end{aligned}$$

- **Example**: One-sided test $(H_A : \mu < \mu_0)$ at a 5% significance level

$$\text{Reject if } t < -1.645$$

- See Wooldridge for more critical values.