Continuous data – 1 group testing

Things we do with data

- Estimation (Descriptive)
  - Summarizing a collection of numbers
  - Identifying the important features of a set of data and extracting useful information
- Testing (Inferential)
  - Drawing conclusions about a population based on information contained in a sample of observations

Things we do with data

- Estimation (Descriptive)
  - Summarizing a collection of numbers
  - Identifying the important features of a set of data and extracting useful information

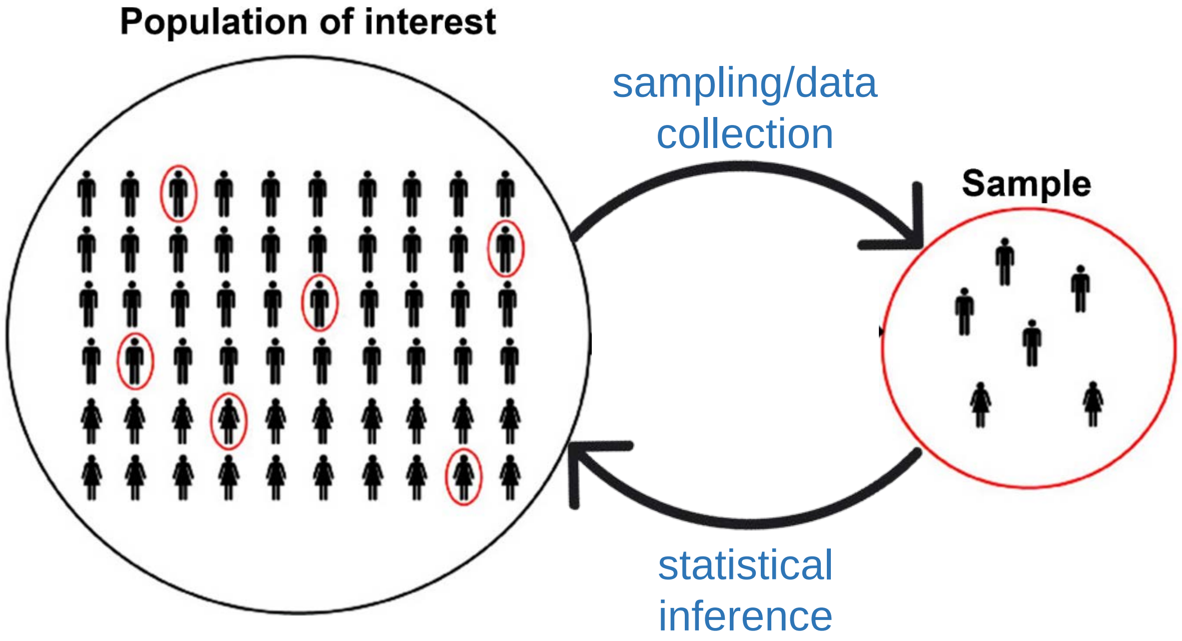
Descriptive summaries
Confidence intervals

- Testing (Inferential)
  - Drawing conclusions about a population based on information contained in a sample of observations

Hypothesis testing

Notation

- μ = population mean
- σ = population standard deviation
- $\bar{x}$ = sample mean
- s = sample standard deviation



Goal: Is the population mean equal to a certain value?

1. State hypotheses
   - Null hypothesis: μ = a certain value
   - Alternative hypothesis: $\mu \neq$ a certain value
2. Collect a sample
3. Assess how much the information in the sample supports the hypotheses
   - P-value
4. Reject or fail to reject the null hypothesis

Hypotheses

population mean

1. State hypotheses

μ_0

H_0

- Null hypothesis: μ = a certain value
- Alternative hypothesis: $\mu \neq$ a certain value

H_A

- Null hypothesis (H_0) is that two things are equal, or there is no change, or nothing happens
- Alternative hypothesis (H_A) is typically what you're trying to show: two things aren't equal, there's a change, something happened, etc.

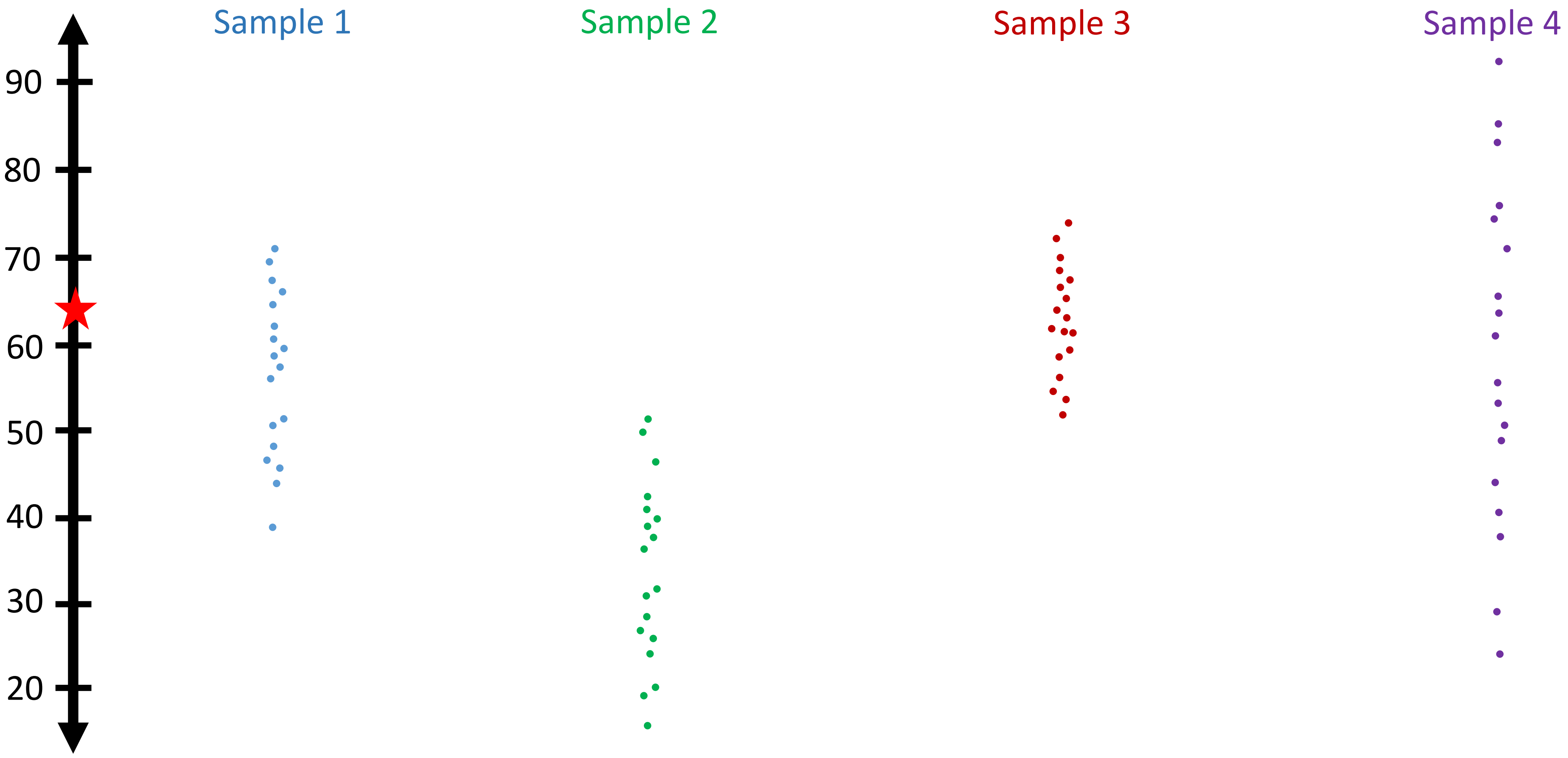
Example: Low Birth Weight SBP

- The average blood pressure among healthy newborn babies is 64/41 (systolic blood pressure = 64 mmHg).
- Question: Do low birth weight infants have a different systolic blood pressure than normal birth weight infants? In other words, is the average systolic blood pressure in low birth weight infants different from 64 mmHg?

H_0:

H_A:

Example: Low Birth Weight SBP



Low Birth Weight Data

- Information on 100 low birth weight infants born in two teaching hospitals in Boston, Massachusetts

Variable	Description
sex	Sex of the baby (Male, Female)
gestage	Gestational age at time of birth (weeks)
length	Length of the baby (cm)
birthwt	Birth weight of the baby (g)
headcirc	Baby's head circumference (cm)
apgar	Apgar score (integers, min=0, max=10). This is a scoring system used for assessing the clinical status of a newborn. 7 or higher is generally considered normal, 4-6 is low, and 3 or below is critically low.

Find the dataset (lowbwt.xlsx) and the full data dictionary (lowbwt Data Dictionary.pdf) in the Data Module on the Canvas site

Example: Low Birth Weight SBP

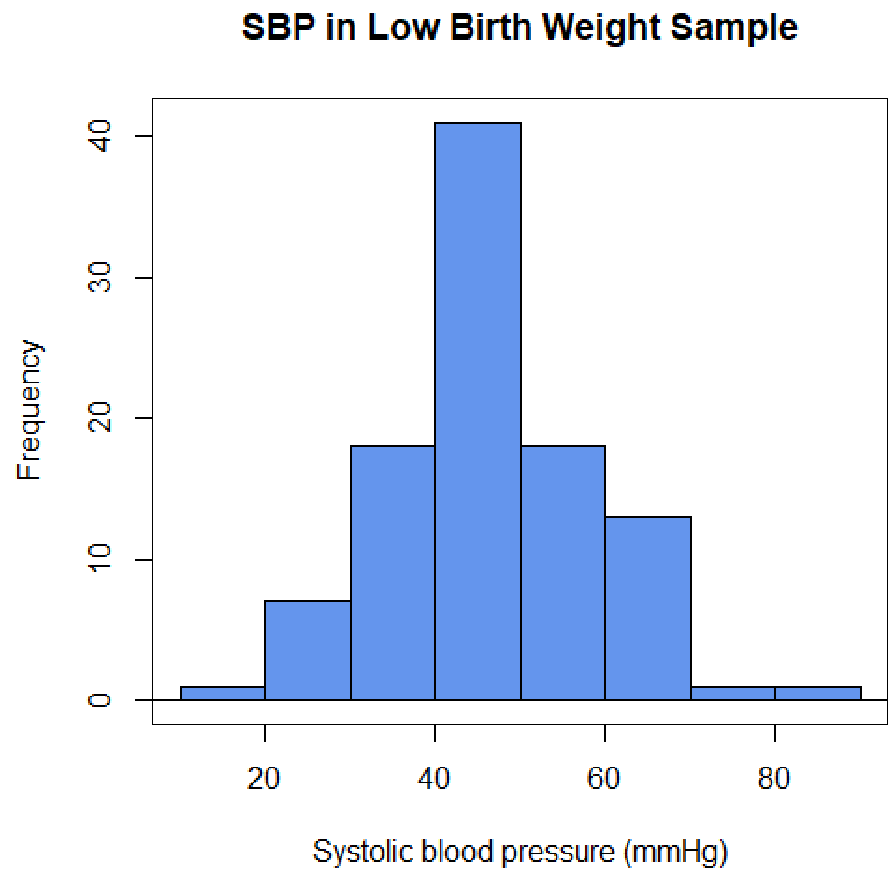
Let μ be the mean SBP of all LBW infants.

H_0: $\mu = 64$ $\qquad$ H_A: $\mu \neq 64$

Descriptive statistics for SBP in sample:

```
 mean       sd   n
47.08 11.40324 100
```

Is this enough evidence to refute H_0?



One-group t Test

- Used when:
  - Variable of interest is continuous
  - Comparing one sample/group to a specific value
  - Population standard deviation (σ) is unknown

- Pre-specify what level of evidence we need to claim that H_0 isn't true (call this α)
- Primary result is the p-value
- If p-value ≤ α, reject H_0

  If p-value > α, fail to reject H_0

P-Value

- Quantifies how much the information in your sample is in agreement with your null hypothesis
  - Higher → our sample is more in agreement with H_0
  - Lower → our sample is less in agreement with H_0
    - Below a certain threshold, we say we have enough evidence to reject H_0

Technical definition: Given the null hypothesis is true, the p-value is the probability of obtaining the results in your sample or more extreme

Significance Level (α)

- A pre-specified threshold that we have to meet for claiming that our sample is showing us evidence against H_0
  - The smaller α is, the more evidence we need to reject H_0
- Think of it as the probability of a false positive that you're willing to have
- Typically $\alpha = 0.05$ is used in scientific literature

Example: Low Birth Weight SBP

H_0: $\mu = 64$ H_A: $\mu \neq 64$ $\alpha = 0.05$

One-group t test for systolic blood pressure:

```
        One Sample t-test

data:  sbp
t = -14.838, df = 99, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 64
95 percent confidence interval:
 44.81735 49.34265
sample estimates:
mean of x
    47.08
```

$\bar{x} = 47.08$
p-value < 0.0001

Since the p-value is less than 0.05, we reject H_0 and conclude that there is sufficient evidence to say that the mean systolic blood pressure in low birth weight infants is not 64 mmHg.

Example: Low Birth Weight Length

- Researchers know that the average length of low birth weight infants in Chicago is 36.4 cm. They are interested in whether the average length of low birth weight infants is different in Boston.
- Use a significance level of 0.10.

H_0:

H_A:

Example: Low Birth Weight Length

$H_0\!: \mu = 36.4 \quad H_A\!: \mu \neq 36.4 \quad \alpha = 0.10$

One-group t test for length:

```
        One Sample t-test

data:  length
t = 1.176, df = 99, p-value = 0.2424
alternative hypothesis: true mean is not equal to 36.4
90 percent confidence interval:
 36.227 37.413
sample estimates:
mean of x
    36.82
```

$\bar{x} = 36.82$
p-value $= 0.24$

Since the p-value is greater than 0.10, we fail to reject H_0 and conclude that there is not sufficient evidence to say that the average length of low birth weight infants in Boston is different from that in Chicago (36.4 cm).

Reject vs. Fail to Reject

- Suppose $\alpha = 0.05$. When p-value ≤ 0.05, we reject H_0. When p-value > 0.05, we fail to reject H_0.
- "Fail to reject" is a weird double negative. Can we just "accept" the null hypothesis? NO!
  - Hypothesis testing does NOT tell you which hypothesis is true. It simply quantifies whether you have enough evidence to reject the null.
  - Think of it like a courtroom: Prosecution tries to show the defendant is guilty *beyond a reasonable doubt* (say, with 95% certainty). Verdict is either "guilty" or "not guilty". The trial cannot prove innocence – it can only say that there's not enough evidence that the defendant is guilty.
  - Here's a link to a blog post that explains the courtroom analogy in more detail

Two-sided vs. One-sided Test

- Pre-specified
- One-sided test is used when you're only interested in detecting a difference in one direction
- Reflected in your alternative hypothesis

  two-sided test H_A: $\mu \neq \mu_0$

  one-sided test H_A: $\mu > \mu_0$ or H_A: $\mu < \mu_0$
- Two-sided is more conservative (and thus more commonly used in scientific literature)

Example: Low Birth Weight Mothers' Age

- Suppose that the average age of mothers giving birth to normal birth weight babies is 28.6 years. We're interested in whether mothers giving birth to low birth weight babies tend to be younger.

H_0:

H_A:

Example: Low Birth Weight Mothers' Age

H_0: $\mu = 28.6$ H_A: $\mu < 28.6$ $\alpha = 0.05$

One-sided one-group t test for mothers' age:

```
        One Sample t-test

data:  momage
t = -1.4541, df = 99, p-value = 0.07453
alternative hypothesis: true mean is less than 28.6
95 percent confidence interval:
     -Inf 28.72339
sample estimates:
mean of x
    27.73
```

$\bar{x} = 27.73$
p-value $= 0.07$

Since the p-value is greater than 0.05, we fail to reject H_0 and conclude that there is not sufficient evidence to say that the average age of mothers giving birth to low birth weight babies is less than 28.6 years.

Relationship between Hypothesis Tests and Confidence Intervals

- Significance level (α) in hypothesis test corresponds to 1-confidence level in confidence interval
  - For example, $\alpha = 0.05$ corresponds to 95% confidence interval
- Reject/fail to reject H_0 in hypothesis test is the same as checking whether μ_0 is in the confidence interval for μ

Relationship between Hypothesis Tests and Confidence Intervals

H_0: $\mu = 64$ $\alpha = 0.05$
H_A: $\mu \neq 64$

One-group t test for systolic blood pressure:

```
        One Sample t-test

data:  sbp
t = -14.838, df = 99, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 64
95 percent confidence interval:
 44.81735 49.34265
sample estimates:
mean of x
    47.08
```

H_0: $\mu = 36.4$ $\alpha = 0.10$
H_A: $\mu \neq 36.4$

One-group t test for length:

```
        One Sample t-test

data:  length
t = 1.176, df = 99, p-value = 0.2424
alternative hypothesis: true mean is not equal to 36.4
90 percent confidence interval:
 36.227 37.413
sample estimates:
mean of x
    36.82
```

Important Points

- Hypothesis testing concepts
    - Hypotheses
    - Significance level
    - P-value
    - Reject/fail to reject
    - One-sided vs. two-sided
- P-value definition and interpretation of hypothesis test
- Relationship between hypothesis tests and confidence intervals