# Chapter 11 Hypotheses Testing for Two Populations



- We now look into comparing two populations.
- This occurs alot in new drug or new methods vs. old drug or old method.
- Chpt 11 covers the following:

  11.3 2-pop matched pairs mean values.
  - same person but before & after
  - husband/wife pairs.

  11.1 2-pop Independent means.
  - group 1 vs group 2 mean value.

  11.2 2-pop Proportions
  - % in favor of Death Penalty in S.Clarita vs. Palmdale

## 11.3 Matched Pairs

Ex Car before and after a tune-up

| Car # | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Before | 35.4 | 35.2 | 31.1 | 31.6 | 26.5 | 23.1 | 25.2 | 32.4 |
| After | 33.8 | 34.3 | 29.6 | 30.9 | 24.9 | 21.8 | 24.3 | 31.3 |
| Difference | 1.6 | 0.9 | 1.5 | 0.5 | 1.6 | 1.3 | 0.9 | 1.1 |

⊛ It turns out that for matched pairs we just do a H.T. on the difference of the parameter being assessed.

In other words,

* **Notations:**

$\bar{d}$ = sample mean of differences

$S_d$ = Stand dev. of the differences for sample

$\mu_d$ = the population mean difference.

* **Assumptions:**

1. $\boxed{\text{SRS}}$ for the single population of matched data {usually before vs after}

2. We need $\boxed{n > 30}$ or the differences need to bell-shaped.

* **Test statistic** (t-table for means)

$$\boxed{t = \frac{\bar{d} - 0}{S_d/\sqrt{n}}}$$

null hypothesis "No differences"

* **Hypothesis:**

$H_0 : \mu_{\text{differences}} = 0$

$H_A : \mu_{\text{differences}} \begin{matrix} < \\ \neq \\ > \end{matrix} 0$ tail

• left one-tail header

$<$

• two tail header

$\neq$

• right one-tail header

$>$

| Assumptions for Inference | And the Conditions That Support or Override Them |
|---|---|

HT 2-pop means, matched pairs

## Proportions ($z$)

- **One sample**

| | |
|---|---|
| 1. Individuals are independent. | 1. SRS and $n < 10\%$ of the population. |
| 2. Sample is sufficiently large. | 2. Successes and failures each $\geq 10$. |

- **Two Groups**

| | |
|---|---|
| 1. Groups are independent. | 1. (Think about how the data were collected.) |
| 2. Data in each group are independent. | 2. Both are SRSs and $n < 10\%$ of populations OR random allocation. |
| 3. Both samples are sufficiently large. | 3. Successes and failures each $\geq 10$ for both groups. |

## Means ($t$)

- **One Sample** (df $= n - 1$)

| | |
|---|---|
| 1. Individuals are independent. | 1. SRS and $n < 10\%$ of the population. |
| 2. Population has a Normal model. | 2. Histogram is unimodal and symmetric.* |

ASSUMPTIONS JUSTIFICATIONS

- **Matched pairs** (df $= n - 1$)

| | |
|---|---|
| 1. Data are matched. | 1. (Think about the design.) |
| 2. Individuals are independent. | 2. SRS and $n < 10\%$ OR random allocation. |
| 3. Population of differences is Normal. | 3. Histogram of differences is unimodal and symmetric.* or n>30 |

- **Two independent samples** (df from technology)

| | |
|---|---|
| 1. Groups are independent. | 1. (Think about the design.) |
| 2. Data in each group are independent. | 2. SRSs and $n < 10\%$ OR random allocation. |
| 3. Both populations are Normal. | 3. Both histograms are unimodal and symmetric.* or both n>30 |

## Distributions/Association ($\chi^2$)

- **Goodness of fit** (df = # of cells $- 1$; one variable, one sample compared with population model)

| | |
|---|---|
| 1. Data are counts. | 1. (Are they?) |
| 2. Data in sample are independent. | 2. SRS and $n < 10\%$ of the population. |
| 3. Sample is sufficiently large. | 3. All expected counts $\geq 5$. |

- **Homogeneity** [df $= (r - 1)(c - 1)$; many groups compared on one variable]

| | |
|---|---|
| 1. Data are counts. | 1. (Are they?) |
| 2. Data in groups are independent. | 2. SRSs and $n < 10\%$ OR random allocation. |
| 3. Groups are sufficiently large. | 3. All expected counts $\geq 5$. |

- **Independence** [df $= (r - 1)(c - 1)$; sample from one population classified on two variables]

| | |
|---|---|
| 1. Data are counts. | 1. (Are they?) |
| 2. Data are independent. | 2. SRSs and $n < 10\%$ of the population. |
| 3. Sample is sufficiently large. | 3. All expected counts $\geq 5$. |

## Regression ($t$, df $= n - 2$)

- **Association of each quantitative variable** ($\beta = 0$?)

| | |
|---|---|
| 1. Form of relationship is linear. | 1. Scatterplot looks approximately linear. |
| 2. Errors are independent. | 2. No apparent pattern in residuals plot. |
| 3. Variability of errors is constant. | 3. Residuals plot has consistent spread. |
| 4. Errors have a Normal model. | 4. Histogram of residuals is approximately unimodal and symmetric, or Normal probability plot reasonably straight.* |

(*less critical as $n$ increases)

**Ex** Will tune-ups improve fuel economy? 

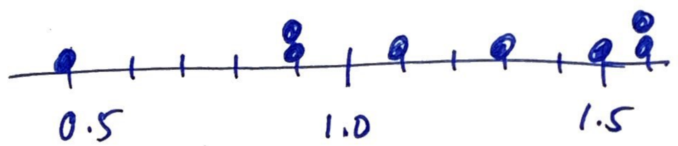
From the example on page one we saw the differences to be: 1.6, 0.9, 1.5, 0.5, 1.6, 1.3, 0.9, 1.1

$\bar{x} = 1.175$  $s_d = 0.396$ ($\sigma_{n-1}$), 0.370 ($\sigma_n$)

Are the improvements just statistical variation OR a statistical significant difference? use $\alpha = 0.01$

**Step ∅:** (a) Type of problem: means, $\sigma$ unkn, matched pair
$\Rightarrow$ t-table

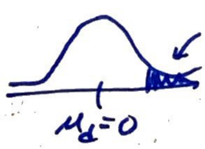
(b) assumptions
- matched data — yes (per car)
- SRS — assumed
- $n < 10\%$ — yes thousands of cars
- Bell shaped? (since $n < 30$)   No, Proceed any ways



dot plot

**Step 1:** hypotheses
(a) $H_0: \mu_d = 0$
$H_A: \mu_d > 0$ {improvement?}

(b) Tail: right tail  (c)



$\alpha$ = area = 0.01

**Step 2:** significance

column
$\boxed{\alpha = 0.01}$

$t_{critical}$ for a one-tail means $\alpha = 0.01$ is $\boxed{t_c = 2.998}$  n=8 source

row
$\boxed{DOF = 8 - 1 = 7}$

**TABLE A-3** *t* Distribution: Critical *t* Values

α =

<, >

C.I., ≠

n=8

| Degrees of Freedom | | | | | |
|---|---|---|---|---|---|
| Area in One Tail | 0.005 | 0.01 | 0.025 | 0.05 | 0.10 |
| Area in Two Tails | 0.01 | 0.02 | 0.05 | 0.10 | 0.20 |
| 1 | 63.657 | 31.821 | 12.706 | 6.314 | 3.078 |
| 2 | 9.925 | 6.965 | 4.303 | 2.920 | 1.886 |
| 3 | 5.841 | 4.541 | 3.182 | 2.353 | 1.638 |
| 4 | 4.604 | 3.747 | 2.776 | 2.132 | 1.533 |
| 5 | 4.032 | 3.365 | 2.571 | 2.015 | 1.476 |
| 6 | 3.707 | 3.143 | 2.447 | 1.943 | 1.440 |
| 7 | 3.499 | 2.998 | 2.365 | 1.895 | 1.415 |
| 8 | 3.355 | 2.896 | 2.306 | 1.860 | 1.397 |
| 9 | 3.250 | 2.821 | 2.262 | 1.833 | 1.383 |
| 10 | 3.169 | 2.764 | 2.228 | 1.812 | 1.372 |
| 11 | 3.106 | 2.718 | 2.201 | 1.796 | 1.363 |
| 12 | 3.055 | 2.681 | 2.179 | 1.782 | 1.356 |
| 13 | 3.012 | 2.650 | 2.160 | 1.771 | 1.350 |
| 14 | 2.977 | 2.624 | 2.145 | 1.761 | 1.345 |
| 15 | 2.947 | 2.602 | 2.131 | 1.753 | 1.341 |
| 16 | 2.921 | 2.583 | 2.120 | 1.746 | 1.337 |
| 17 | 2.898 | 2.567 | 2.110 | 1.740 | 1.333 |
| 18 | 2.878 | 2.552 | 2.101 | 1.734 | 1.330 |
| 19 | 2.861 | 2.539 | 2.093 | 1.729 | 1.328 |
| 20 | 2.845 | 2.528 | 2.086 | 1.725 | 1.325 |
| 21 | 2.831 | 2.518 | 2.080 | 1.721 | 1.323 |
| 22 | 2.819 | 2.508 | 2.074 | 1.717 | 1.321 |
| 23 | 2.807 | 2.500 | 2.069 | 1.714 | 1.319 |
| 24 | 2.797 | 2.492 | 2.064 | 1.711 | 1.318 |
| 25 | 2.787 | 2.485 | 2.060 | 1.708 | 1.316 |
| 26 | 2.779 | 2.479 | 2.056 | 1.706 | 1.315 |
| 27 | 2.771 | 2.473 | 2.052 | 1.703 | 1.314 |
| 28 | 2.763 | 2.467 | 2.048 | 1.701 | 1.313 |
| 29 | 2.756 | 2.462 | 2.045 | 1.699 | 1.311 |
| 30 | 2.750 | 2.457 | 2.042 | 1.697 | 1.310 |
| 31 | 2.744 | 2.453 | 2.040 | 1.696 | 1.309 |
| 32 | 2.738 | 2.449 | 2.037 | 1.694 | 1.309 |
| 34 | 2.728 | 2.441 | 2.032 | 1.691 | 1.307 |
| 36 | 2.719 | 2.434 | 2.028 | 1.688 | 1.306 |
| 38 | 2.712 | 2.429 | 2.024 | 1.686 | 1.304 |
| 40 | 2.704 | 2.423 | 2.021 | 1.684 | 1.303 |
| 45 | 2.690 | 2.412 | 2.014 | 1.679 | 1.301 |
| 50 | 2.678 | 2.403 | 2.009 | 1.676 | 1.299 |
| 55 | 2.668 | 2.396 | 2.004 | 1.673 | 1.297 |
| 60 | 2.660 | 2.390 | 2.000 | 1.671 | 1.296 |
| 65 | 2.654 | 2.385 | 1.997 | 1.669 | 1.295 |
| 70 | 2.648 | 2.381 | 1.994 | 1.667 | 1.294 |
| 75 | 2.643 | 2.377 | 1.992 | 1.665 | 1.293 |
| 80 | 2.639 | 2.374 | 1.990 | 1.664 | 1.292 |
| 90 | 2.632 | 2.368 | 1.987 | 1.662 | 1.291 |
| 100 | 2.626 | 2.364 | 1.984 | 1.660 | 1.290 |
| 200 | 2.601 | 2.345 | 1.972 | 1.653 | 1.286 |
| 300 | 2.592 | 2.339 | 1.968 | 1.650 | 1.284 |
| 400 | 2.588 | 2.336 | 1.966 | 1.649 | 1.284 |
| 500 | 2.586 | 2.334 | 1.965 | 1.648 | 1.283 |
| 750 | 2.582 | 2.331 | 1.963 | 1.647 | 1.283 |
| 1000 | 2.581 | 2.330 | 1.962 | 1.646 | 1.282 |
| 2000 | 2.578 | 2.328 | 1.961 | 1.646 | 1.282 |
| Large | 2.576 | 2.326 | 1.960 | 1.645 | 1.282 |

EX (cont.) 

4

**Step 3: Test statistic**

(a) SE : $SE = s_d/\sqrt{n}$ ← need std. dev. of the differences! $\sigma_{n-1}$ on calc.

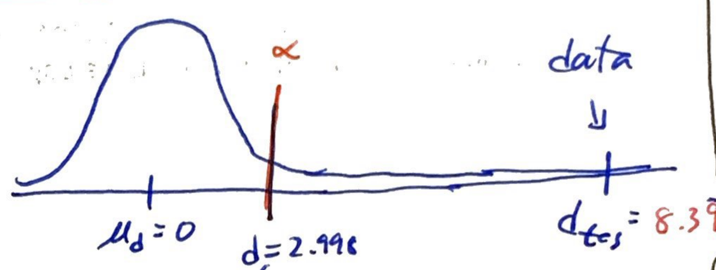
$$= \frac{0.396}{\sqrt{8}}$$

$$= \underline{\underline{0.140}}$$

(b)

$$t_{test} = \frac{\text{sample} - \text{null}}{SE} = \frac{1.175 - 0}{0.140} = \underline{\underline{8.39}}$$

**Step 4:** p-value (—OR— critical vs test)



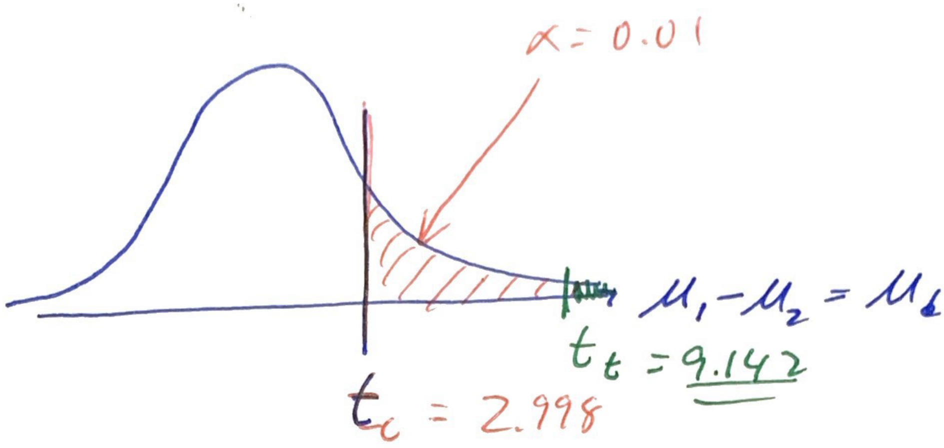

**Step 5: Interpret results**

Since $d_{test} \gg d_{crit}$ we accept the claim that fuel economies improve after a tune up.

**Step 6:** "It was discovered that their is significant improvement in a car's fuel economy after a tune-up"

**Step 7:** statdisk:

STEP 6: The critical value method




Conclusion: since $t_{test}$ is further away from 0 than $t_{crit}$ we reject the null hypothesis that the difference in fuel economies before and after a tune-up is negligible ("0")

So tune-ups improve fuel economy.

STEP 7: Statdisk → analysis → hyp Test → mean matched

- populate Sample Editor with Col 1 = after, Col 2 = before

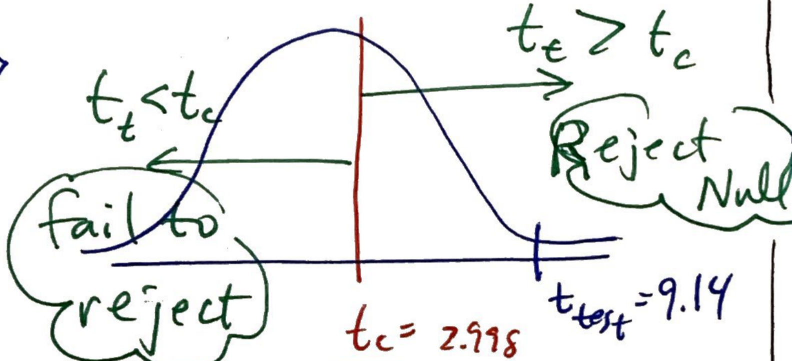
[ > ] right tail

α = [0.01]

- pop 1 → [Col 1]
- pop 2 → [Col 2]

[Evaluate]

pvalue = 0.00002




[plot]