Statistical Methods for Plant Biology

PBIO 3150/5150

Anirudh V. S. Ruhil

March 8, 2016

The Voinovich School of Leadership and Public Affairs

Table of Contents

Paired $t-tests$

Paired $t - tests$

Definition

In a paired design both treatments are applied to every sampled unit

Example

- Patients before and after hospitalization
- Students before and after tutoring
- Blackbirds before and after testosterone implants
- Cholesterol before and after diet-plus-exercise-regimen
- Comparing 30-day readmission rates of patients not given pre-operative therapy with "matched" patients given pre-operative therapy

Assumptions

1. Random samples
2. Differences are approximately Normally distributed (... but Y_1 and Y_2 can follow any distribution)

The Test Statistic and the Testing Protocol

$$\begin{aligned} d_i &= Y_1 - Y_2 \\ \bar{d} &= \frac{\sum d_i}{n} \\ s_d^2 &= \frac{\sum (d_i - \bar{d})^2}{n-1} \\ s_d &= \sqrt{\frac{\sum (d_i - \bar{d})^2}{n-1}} \end{aligned}$$

Test Statistic: $t = \frac{\bar{d} - \mu_d}{s_d/\sqrt{n}}; df = n-1$

Interval Estimate: $\bar{d} \pm t_{\alpha/2}\left(\frac{s_d}{\sqrt{n}}\right)$

Specify Hypotheses:

$H_0 : \mu_d = 0; H_A : \mu_d \neq 0$

$H_0 : \mu_d \leq 0; H_A : \mu_d > 0$

$H_0 : \mu_d \geq 0; H_A : \mu_d < 0$

Set α

Conduct the Test

Reject H_0 if $P-value \leq \alpha$; Do not reject H_0 otherwise

An Example

Child	Y_1	Y_2	d_i
1	6.0	5.4	0.6
2	5.0	5.2	-0.2
3	7.0	6.5	0.5
4	6.2	5.9	0.3
5	6.0	6.0	0.0
6	6.4	5.8	0.6

$\bar{d} = 0.30$; $s_d = 0.335$

$t = 2.1958$, df = 5, $P-value = 0.07952$

95% CI: $0.3 \pm 0.35 = (-0.0512, 0.6512)$

Do not reject H_0; Data do not suggest that the method used influences learning

Testosterone and Blackbirds

In many species, males are more likely to attract females if the males have high testosterone levels. Are males with high testosterone paying a cost for this extra mating success in other ways? One hypothesis is that males with high testosterone might be less able to fight off disease – that is, their high levels of testosterone might reduce their immunocompetence. To test this hypothesis researchers artificially increased the testosterone levels of 13 male red-winged blackbirds by implanting a small permeable tube filled with testosterone. They measured immunocompetence as the rate of antibody production in response to a nonpathogenic antigen in each bird's blood serum both before and after the implant. The antibody production rates were measured optically, in units of log 10^{-3} optical density per minute $(ln[mOD/min])$

Testosterone and Blackbirds

before	after	log.before	log.after	diff.in.logs	diff
105	85	4.65	4.44	0.21	20
50	74	3.91	4.30	-0.39	-24
136	145	4.91	4.98	-0.07	-9
90	86	4.50	4.45	0.05	4
122	148	4.80	5.00	-0.20	-26
132	148	4.88	5.00	-0.12	-16
131	150	4.88	5.01	-0.13	-19
119	142	4.78	4.96	-0.18	-23
145	151	4.98	5.02	-0.04	-6
130	113	4.87	4.73	0.14	17
116	118	4.75	4.77	-0.02	-2
110	99	4.70	4.60	0.10	11
138	150	4.93	5.01	-0.08	-12

H_0: Mean change in antibody production after testosterone implants was zero ($\mu_d = 0$)

H_A: Mean change in antibody production after testosterone implants was not zero ($\mu_d \neq 0$)

Let $\alpha = 0.05$

```
> t.test(Blackbirds$log.before, Blackbirds$log.after, paired=TRUE, conf.level=0.95, alternative
      ="two.sided")

	Paired t-test

data:  x and Blackbirds$log.after
t = -1.2714, df = 12, p-value = 0.2277
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.15238464 0.04007695
sample estimates:
mean of the differences
           -0.05615385
```

Because the $P-value$ is not $\leq \alpha$ we do not Reject H_0; the data suggest that mean change in antibody production after testosterone implants was zero.

Two-Sample Comparison of Means

Two-Sample Designs

Two-Tailed Hypotheses

H_0: $\mu_1 = \mu_2$; H_A: $\mu_1 \neq \mu_2$
These can be rewritten as H_0: $\mu_1 - \mu_2 = 0$; H_A: $\mu_1 - \mu_2 \neq 0$

One-Tailed Hypotheses

H_0: $\mu_1 \leq \mu_2$; H_A: $\mu_1 > \mu_2$
These can be rewritten as H_0: $\mu_1 - \mu_2 \leq 0$; H_A: $\mu_1 - \mu_2 > 0$

H_0: $\mu_1 \geq \mu_2$; H_A: $\mu_1 < \mu_2$
These can be rewritten as H_0: $\mu_1 - \mu_2 \geq 0$; H_A: $\mu_1 - \mu_2 < 0$

Test Statistic

$$t = \frac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{SE_{\bar{Y}_1 - \bar{Y}_2}}$$

The pooled sample variance: $s_p^2 = \dfrac{df_1 s_1^2 + df_2 s_2^2}{df_1 + df_2}$; $df_1 = n_1 - 1; df_2 = n_2 - 1$

The Test Statistic

The standard error $SE_{\bar{Y}_1 - \bar{Y}_2}$ and df are calculated in one of two ways:

1. Assuming equal population variances

$$SE_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \, ; df = n_1 + n_2 - 2$$

2. Assuming unequal population variances

$$SE_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

approximate $df = \dfrac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2} \right)^2}{\left[\dfrac{\left(\frac{s_1^2}{n_1} \right)^2}{df_1} + \dfrac{\left(\frac{s_2^2}{n_2} \right)^2}{df_2} \right]}$ … rounded down to nearest integer

Assumptions and Rules-of-thumb

Assumptions:

1. Random samples
2. Variables are from normally distributed Populations (formally tested)
3. Variables have equal variances in the Population (formally tested)

Rules-of-thumb:

- Draw larger samples if you suspect the Population(s) may be skewed
- Go with *equal variances* if both the following are met:
    1. Assumption theoretically justified, standard deviations fairly close
    2. $n_1 \geq 30$ and $n_2 \geq 30$
- Go with *unequal variances* if both the following are met:
    1. One standard deviation is at least twice the other standard deviation
    2. $n_1 < 30$ or $n_2 < 30$

Horned Lizards and the Loggerhead Shrike

The horned lizard has many unusual features, including its ability to squirt blood from its eyes. Herpetologists recently tested the idea that long spikes help protect horned lizards from being eaten. They focused on the remains of 30 horned lizards killed by shrikes (which impales its victims on thorns or barbed wires to save for later eating). Horn length was measured for these 30 victims.

As a comparison group they measured horn lengths for 154 horned lizards still alive and well. They then compared the mean horn lengths of the two groups.

Horned Lizards and the Loggerhead Shrike

H_0: Mean horn lengths do not differ between lizards killed by shrikes and lizards still alive $(\mu_1 = \mu_2)$
H_A: Mean horn lengths differ between lizards killed by shrikes and lizards still alive $(\mu_1 \neq \mu_2)$
Set $\alpha = 0.05$

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{SE_{\bar{Y}_1 - \bar{Y}_2}}$$

```
> t.test(HornedLizards$horn.length ~ HornedLizards$group, conf.level=0.95, var.equal=TRUE)

	Two Sample t-test

data:  HornedLizards$horn.length by HornedLizards$group
t = -4.3494, df = 182, p-value = 2.27e-05
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -3.335402 -1.253602
sample estimates:
mean in group killed mean in group living 
            21.98667             24.28117 
```

Cichlids and their Preferences

The astonishing diversity of cichlid fishes of Lake Victoria is maintained by the preferences of females for males of their own species. To understand how the species arose in the first place, it is important to know the genetic basis of this preference in females.

Researchers crosses two species of cichlids, *Pundamilia pundamilia* and *P.nyererei*, and raised the “F1 hybrids” to adulthood. The measured degree of preference by the female F1 fish for *P.pundamilia* males over *P.nyererei* males. They then crossed the F1 hybrids with each other to produce a second generation of hybrids (the F2), which they also raised to adulthood and measured the same index of female preference.

If a small number of genes are important in determining the preference, then the variance of the preference index will differ between the these two generations (it will be highest in F2 hybrids). The researchers measured preference in 20 F1 and 33 F2 individuals.

Does the mean preference differ between F1 and F2 hybrids?

H_0: The mean preference is the same for the two groups $(\mu_1 - \mu_2 = 0)$
H_A: The mean preference is not the same for the two groups $(\mu_1 - \mu_2 \neq 0)$
$\alpha = 0.05$

Using R to conduct the test we obtain:

$t = -0.0874, df = 51, p - value = 0.9307$ ... assuming equal variances, and

$t = -0.1047, df = 46.042, p - value = 0.9171$ ... assuming unequal variances

Regardless, we fail to reject H_0 under either assumption; the data provide insufficient evidence to conclude that the mean preference is not the same for the two groups.

Some Cautions

Correct Sampling Units

One of the greatest threats to biodiversity is the introduction of alien species from outside their natural range. These introduced species often have fewer predators or parasites in the new area so they can increase in number and outcompete native species. The brook trout, for example, is a species native to eastern North America that has been introduced into streams in the West for sport fishing.

Biologists followed the survivorship of a native species, the chinook salmon, in a series of 12 streams that either had brook trout or did not. Their goal was to determine whether the presence of brook trout affected the survivorship of the salmon. In each stream, they released a number of tagged juvenile chinook and then recorded whether or not each chinook survived over one year.

trout	survived	released	$\hat{p}_{surviving}$
present	166	820	0.20
absent	180	467	0.39
present	136	960	0.14
present	153	700	0.22
absent	178	959	0.19
present	103	545	0.19
absent	326	1029	0.32
present	173	769	0.22
absent	7	27	0.26
absent	120	998	0.12
absent	135	936	0.14
present	188	1001	0.19

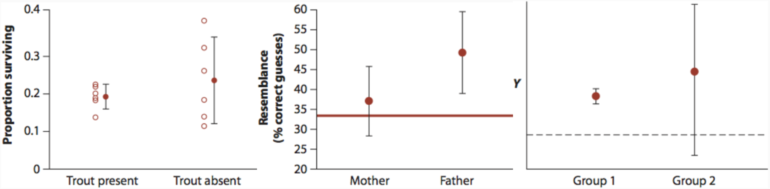
Tempting to create 2×2 table of Trout present/absent and Chinook survived/died, and then carry out ta χ^2 test. This would yield a significant p-value, suggesting that there is an association between the presence of brook trout and survival.

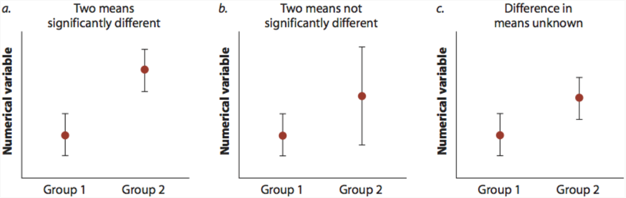
... but chinook are clustered in streams and thus not sampled independently of one another within a stream

Appropriate analysis would be a two-sample $t-test$, and this suggests there is no difference in the mean proportion of chinook surviving in streams with and without brook trout

Overlapping Confidence Intervals

- Never compare individual group means to the null hypothesized value; always compare groups to each other
- Interpret overlapping confidence intervals with care





Comparing Variances

The $F-test$ for Equality of Variances

- Assumes normally distributed populations (hence sensitive to departures)
- If X_1 and X_2 are two independent random variables distributed as χ^2 with df_1 and df_2, respectively,then the ratio $\frac{\frac{X_1}{df_1}}{\frac{X_2}{df_2}}$ follows the F distribution with df_1 in the numerator and df_2 in the denominator
- Hypotheses: $H_0 : \sigma_1^2 = \sigma_2^2; H_A : \sigma_1^2 \neq \sigma_2^2$
- Test Statistic: $F = \frac{s_1^2}{s_2^2}; F \sim F_{\alpha/2, df_1, df_2}$
- Note: s_1^2 is the larger sample variance

Levene's Test for Homogeneity of Variances

- Assumes roughly symmetric frequency distributions within all groups
- Robust to violations of assumption
- Can be used with 2 or more groups
- $H_0 : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \cdots \sigma_k^2$
  H_A : For at least one pair of (i, j) we have $\sigma_i^2 \neq \sigma_j^2$
- Test Statistic: $W = \dfrac{(N-k)\sum_{i=1}^{k} n_i \left(\bar{Z}_i - \bar{Z}\right)^2}{(k-1)\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(Z_{ij} - \bar{Z}_i\right)^2}$
- $Z_{ij} = |Y_{ij} - \bar{Y}_i|$; $\bar{Z}_i$ is the mean for all Y in the i^{th} group; $\bar{Z}$ is the mean for all Y in the study; k is the number of groups in the study; and n_i is the sample size for group i
- If you opt for the more robust version that uses the Median, then, $Z_{ij} = |Y_{ij} - \tilde{Y}_i|$ where $\tilde{Y}_i$ is the median of the i^{th} group
- $W \sim F_{\alpha,k-1,n-k}$

Problem # 9

(a) I'll use the $F-test$ for equality of variances ... $H_0: \sigma_1^2 = \sigma_2^2; H_A: \sigma_1^2 \neq \sigma_2^2$

```
> var.test(preference ~ genotype, data=Cichlids, conf.level=0.95, alternative="two.sided", ratio=1)
       F test to compare two variances
data:  preference by genotype
F = 0.1649, num df = 19, denom df = 32, p-value = 0.0001266
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.0755988 0.3922485
sample estimates:
ratio of variances
        0.1648966
```

(b) Default to assumption of unequal variances (both conditions are met)
H_0: Mean preference index is the same for both groups $(\mu_1 = \mu_2)$
H_A: Mean preference index is not the same for both groups $(\mu_1 \neq \mu_2)$

```
> t.test(preference ~ genotype, data=Cichlids, conf.level=0.95, paired=FALSE, var.equal=FALSE)
       Welch Two Sample t-test
data:  preference by genotype
t = -0.1047, df = 46.042, p-value = 0.9171
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.06577309 0.05927006
sample estimates:
mean in group F1 mean in group F2
    0.004900000     0.008151515
```

Testing Options and the Protocol

- Data are coming from a "paired" design – Use the two-sample $t-test$ with paired=TRUE
- Data are coming from two "unpaired" groups – Use the two-sample $t-test$ with
  - the assumption of equal variances if $n_1 \geq 30$ and $n_2 \geq 30$ and $s_1 \approx s_2$
  - the assumption of unequal variances if $n_1 < 30$ or $n_2 < 30$ and $s_i \geq 2\left(s_j\right)$
  - Use the $F-test$ to test for equality of variances if the distribution seems normal for each group
  - Use Levene's test for homogeneity of variances if the assumption of normality is not supported