Why?

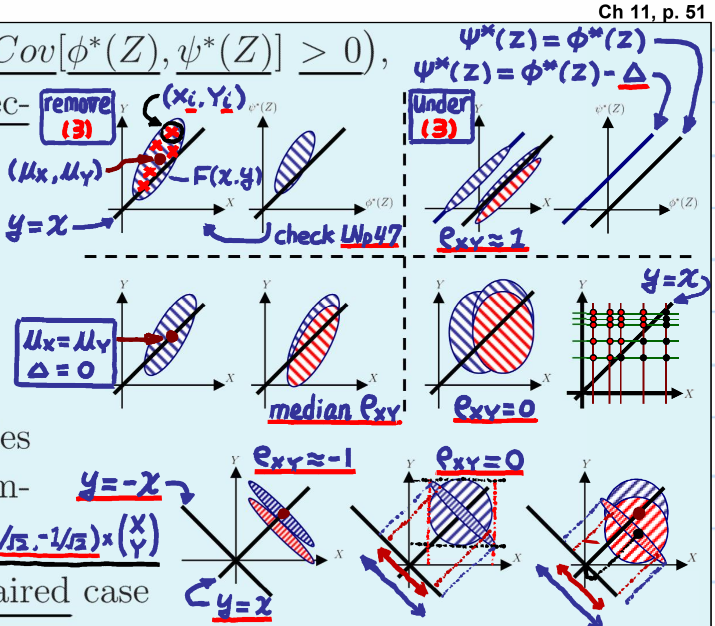
- If $\rho_{XY} > 0$ ($\Leftrightarrow \sigma_{XY} > 0 \Leftrightarrow Cov[\phi^*(Z), \psi^*(Z)] > 0$), then paired sample is more effective than independent samples.

Under (3)

- When $\psi^*(Z) = \phi^*(Z) - \Delta$,

$$Cov[\phi^*(Z), \psi^*(Z)] = Cov[\phi^*(Z), \phi^*(Z) - \Delta] = Var[\phi^*(Z)] > 0.$$

11/13



- **Q**: Why are independent samples more effective than paired samples when $\sigma_{XY} < 0$?

$X - Y = \sqrt{2} \times (1/\sqrt{2}, -1/\sqrt{2}) \times \binom{X}{Y}$

- If $\sigma^2_X = \sigma^2_Y = \sigma^2$, then in the paired case

$$\sigma^2_{\bar{D}} = Var(\bar{D}) = [2\,\sigma^2(1 - \rho_{XY})]/n \quad \text{and} \quad \sigma^2_{\bar{X}-\bar{Y}} = Var(\bar{X} - \bar{Y}) = 2\,\sigma^2/n$$

in the unpaired case. The relative efficiency is $\sigma^2_{\bar{D}}/\sigma^2_{\bar{X}-\bar{Y}} = 1 - \rho_{XY}$.

- If $\rho_{XY} = 0.5$, a paired design with n pairs of subjects yields the same precision as an unpaired design with $2n$ subjects per treatment.

paired
n X_i's
n Y_i's
2-indep.
$2n$ X_i's
$2n$ Y_j's

σ^2_δ / σ^2_ϵ under (3)

- From now on, the analyses of paired data are based on

$$D_i = X_i - Y_i, \quad i = 1, \ldots, n.$$

- Statistical modeling for D_i's: $D_1, \ldots, D_n \sim$ i.i.d. F ⇐ one-sample model

Example 7 (Effect of cigarette smoking on platelet aggregation, Levine, 1973)

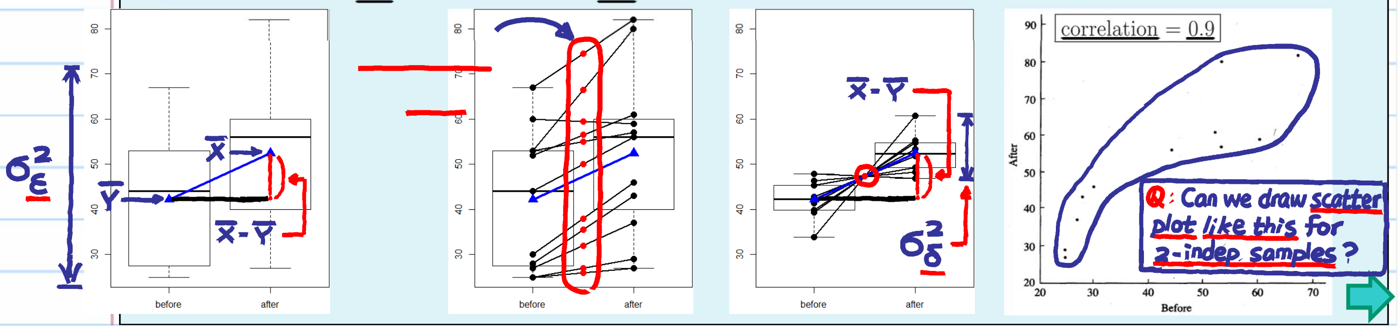
- Blood samples were drawn from 11 individuals before and after they smoked a cigarette to measure the extent to which the blood platlets aggregated.
- data: maximum percentage of all platelets that aggregated after being exposed to a stimulus.

block (LNp.46) →	1	2	3	4	5	6	7	8	9	10	11
before (Y)	25	25	27	44	30	67	53	53	52	60	28
after (X)	27	29	37	56	46	82	57	80	61	59	43
difference (D)	2	4	10	12	16	15	4	27	9	−1	15

paired

Q: Can we take this difference on the X's, Y's data from 2-indep. samples?

- **Q**: Do the differences D_i's indicate a clear pattern of $\Delta = \mu_X - \mu_Y \neq 0$?
- The two-sample (unpaired) t-test for the before and after data gives a p-value = 0.1721 ⇒ the null $H_0 : \mu_X = \mu_Y$ is not rejected under $\alpha = 0.1$ ($s_p^2 = 289.34$).
  - **Q**: Why did the 2-sample t-test not reject H_0 when the differences showed such a clear pattern of $\mu_X > \mu_Y$?
  - Note that in 2-sample t-test, the test statistic is $|T| = \dfrac{|\bar{X} - \bar{Y}|}{s_p\sqrt{\frac{1}{n} + \frac{1}{m}}}$, where s_p^2 estimates σ^2_ϵ, rather than σ^2_δ.



Q: Can we draw scatter plot like this for 2-indep samples?

- Figure 11.7 (textbook, p.447) plots after-values vs. before-values. They are positively correlated with a sample correlation coefficient 0.9. Pairing was a natural and effective experimental design in this case: relative efficiency = 0.1.

- Methods based on normality assumptions

- **Recall**. $D_i = X_i - Y_i$, $i = 1, \ldots, n$, and $D_1, \ldots D_n \sim$ i.i.d. F. ← 1-sample model
- Assume that F is Normal.
- Thus, the statistical model for D_i's is

$$D_1, \ldots D_n \sim \text{i.i.d. } N(\mu_D, \sigma_D^2), \quad (\triangle)$$

← 1-sample normal model

where $\mu_D = \mu_X - \mu_Y$.

  - This model contains two parameters: μ_D ($\in \mathbb{R}$) and σ_D^2 (> 0).
  - Under this model, we can only examine whether there exists "difference" between the means of the two paired samples, i.e.,

$$\mu_D = \mu_X - \mu_Y = 0 \Rightarrow \text{no difference or no effect}$$

cannot examine $\sigma_X^2 \overset{?}{=} \sigma_Y^2$

Theorem 18 (test and confidence interval for μ_D, 1-sample normal model, paired data)

Consider the model ($\triangle$).

- **Recall** (Review 1 in LNp.6-7).
  - $\overline{D} = \frac{1}{n}\sum_{i=1}^{n} D_i \xrightarrow{e} \mu_D$, and $\underbrace{\overline{D}}_{\text{statistic}} \sim N(\mu_D, \sigma_D^2/n) \underset{\text{standardization}}{\Rightarrow} \underbrace{\sqrt{n}(\overline{D} - \mu_D)/\sigma_D}_{\text{pivotal quantity for } \mu_D \text{ if } \sigma_D \text{ known}} \sim N(0,1)$
  - $s_D^2 = \frac{\sum_{i=1}^{n}(D_i - \overline{D})^2}{n-1} \xrightarrow{e} \sigma_D^2$, and $\overbrace{(n-1)s_D^2}^{\text{statistic}} \sim \sigma_D^2 \chi_{n-1}^2 \Rightarrow \underbrace{(n-1)s_D^2/\sigma_D^2}_{\text{pivotal quantity for } \sigma_D^2} \sim \chi_{n-1}^2$
  - $\overline{D}$ and s_D^2 are independent

- Pivotal quantity for μ_D
  - σ_D known: $$Q_{Z,\mu_D} \equiv \frac{\overline{D} - \mu_D}{\sigma_{\overline{D}}} = \frac{\overline{D} - \mu_D}{\sigma_D/\sqrt{n}} = \frac{\sqrt{n}(\overline{D} - \mu_D)}{\sigma_D} \sim N(0,1)$$
  - $\underset{\uparrow \text{parameter}}{\sigma_D}$ unknown: $$Q_{T,\mu_D} \equiv \frac{\overline{D} - \mu_D}{s_{\overline{D}}} = \frac{\overline{D} - \mu_D}{s_D/\sqrt{n}} = \frac{\overbrace{\sqrt{n}(\overline{D} - \mu_D)/\sigma_D}^{\sim N(0,1)}}{\sqrt{\underbrace{\frac{(n-1)s_D^2/\sigma_D^2}{n-1}}_{\sim \chi_{n-1}^2}}} \sim t_{n-1}$$

cf. ($\sigma_{\overline{D}} \leftrightarrow \sigma_D/\sqrt{n}$), cf. ($s_{\overline{D}} \leftrightarrow s_D/\sqrt{n}$), indep.

μ_D: fixed, data: changed

- Test the null and alternative hypotheses at significance level α:

$$H_0 : \mu_D = \mu_{D,0} \quad \text{vs.} \quad H_A : \mu_D \neq \mu_{D,0}$$

where $\mu_{D,0}$ is a known constant.

(**Note**. if $\mu_{D,0} = 0$, this is equivalent to $H_0 : \mu_X = \mu_Y$ vs. $H_A : \mu_X \neq \mu_Y$.)

The two tests are likelihood ratio tests (exercise)

  - σ_D known: reject H_0 if

$$|Z \equiv \frac{\overline{D} - \mu_{D,0}}{\sigma_{\overline{D}}}| > z(\alpha/2) \Leftrightarrow |\overline{D} - \mu_{D,0}| > z(\alpha/2)\,\sigma_{\overline{D}} = z(\alpha/2)\frac{\sigma_D}{\sqrt{n}}$$

  - σ_D unknown: reject H_0 if

$$|T \equiv \frac{\overline{D} - \mu_{D,0}}{s_{\overline{D}}}| > t_{n-1}(\alpha/2) \Leftrightarrow |\overline{D} - \mu_{D,0}| > t_{n-1}(\alpha/2)\,s_{\overline{D}} = t_{n-1}(\alpha/2)\frac{s_D}{\sqrt{n}}$$

estimate difference between μ_D & $\mu_{D,0}$

scale

∵ 2-sided tests

paired samples → one-sample Z-test, one-sample T-test

- A $100(1-\alpha)\%$ confidence interval for μ_D is (data fixed, μ_D: changed)
  - σ_D known: $\overline{D} \pm z(\alpha/2) \times \sigma_{\overline{D}} \Leftrightarrow \overline{D} \pm z(\alpha/2) \times (\sigma_D/\sqrt{n})$
  - σ_D unknown: $\overline{D} \pm t_{n-1}(\alpha/2) \times s_{\overline{D}} \Leftrightarrow \overline{D} \pm t_{n-1}(\alpha/2) \times (s_D/\sqrt{n})$

Example 8 (Effect of smoking, t-test for paired data, cont. Ex.7 In LNp.52)

- $n = 11$, $D_i = \text{after}_i - \text{before}_i$ estimate of population variance σ_D^2
- $\overline{D} = 10.27$, $s_{\overline{D}} = 2.405$ $(\Rightarrow s_D^2 = 11 \times 2.405^2 = 63.62$ cf.

  Same as $\overline{X}-\overline{Y}$ in 2-sample T-test (Ex 7)

  $$\Rightarrow 63.62/2 = 31.81 \overset{e}{\to} \sigma_\delta^2 \overset{\text{cf.}}{\longleftrightarrow} s_p^2 = 289.34 \overset{e}{\to} \sigma_\epsilon^2 \text{ in Ex.7})$$

  cf. include variation of Zi's

If $\sigma_{\delta_1}^2 = \sigma_{\delta_2}^2 = \sigma_\delta^2$, then (LNp.49) $Var(D_i) \doteq 2\sigma_\delta^2$

- A 90% confidence interval for μ_D is α = 0.1

  $$\overline{D} \pm t_{10}(0.05)\, s_{\overline{D}} = 10.27 \pm 1.812 \times 2.40 = (5.9, 14.6), \rightarrow \text{reject } \Delta = \mu_X - \mu_Y = 0$$

  which does not contain zero ($\overset{\text{cf.}}{\longleftrightarrow}$ H_0 not rejected in Ex.7 using 2-sample t-test)
- The (one-sample) t-statistic is $T = (10.27-0)/2.40 = 4.28 > t_{10}(0.005) = 3.169$. The p-value of a two-sided test is less than 0.01. There is little doubt that smoking increases platelet aggregation. α

Note 9 (Some notes about one-sample t-test when normality assumption does not hold)

- Consider the model: $D_1, \ldots D_n \sim$ i.i.d. F, where F can be *any* continuous distributions with *finite* variance.
  - By CLT and LLN, when $n \to \infty$ (sample size is large), → can use 1-sample T-test

    Check Thm12 in LN, CH7, p.29 →

    $$\overline{D} \overset{D}{\approx} N(\mu_D, \sigma_D^2/n) \quad \text{and} \quad s_D^2 \xrightarrow{P} \sigma_D^2.$$
  - Thus, by Slutsky's Thm,

    $$Q_{T,\mu_D} = \frac{(\overline{D}-\mu_D)/(\cancel{\sigma_D}/\sqrt{n})}{\sqrt{s_D^2/\cancel{\sigma_D^2}}} \overset{D}{\approx} N(0,1)$$

    $\overset{D}{\approx} N(0,1)$ $\xrightarrow{P} 1$ not a function of σ_D^2

    and t_{n-1} tends to $N(0,1)$ as $n \to \infty$.
- **Q**: What if the sample size n is small or population variance $= \infty$? cf.