

Homework # 1 Solutions

Biostatistics 210

0. PURPOSE

This assignment was intended to give you a framework for comparing a series of analyses using the concepts of validity (tests with correct size, unbiasedness) and efficiency (relative efficiency and power of tests). The setting in which you did it was a series of possible analyses for a simple clinical trial with repeated measures.

The data setting was a clinical trial with 100 individuals randomly assigned between placebo and a drug. The idea was to examine the effect of the drug on measures at week 24 given that there was a baseline measure.

The model that generated the data followed

	Mean Global Deficit Score	
	Baseline	Week 24
Placebo	α	$\alpha + \beta$
Minocycline	α	$\alpha + \beta$

The model has important implications.

First, it says the model for the mean value at week 24 is

$$1. \text{mean(gds2)} = \alpha + \beta_0 + (\beta_1 * rx)$$

where rx is a binary variable (1=minocycline, 0=placebo). This is exactly the setting for the t-test -- a setting with a mean between two groups. The t-test identifies the differences in means which is just β_1 .

$$2. \text{mean(diff)} = \beta_0 + \beta_1 * rx$$

for the variable "diff" the mean is 0 in the placebo group and β_1 in the placebo group

$$3. E(\text{gds2} | \text{gds1}) = \alpha + b_0 \text{gds1} + (\beta_1 * rx)$$

here the formula is not obvious and is a consequence of the way I generated the data

Then later we transformed the data to have gds be the outcome in a repeated measures model and we created the variables timepoint (0=baseline, 1=week 24) and a product variable

$$4. \text{mean(gds)} = \alpha + \beta_0 \text{ timepoint} + \beta_1 * \text{inter}$$

based on this. The mean GDS is equal in the two groups at baseline (as expected in a clinical trial) and value is set to be α . The mean GDS is the placebo group at week 24 is $\alpha + \beta_0$ in the placebo group and is $\alpha + \beta_0 + \beta_1$ in the minocycline group at week 24.

3. ANALYSIS

We examined four different ways of estimating this difference

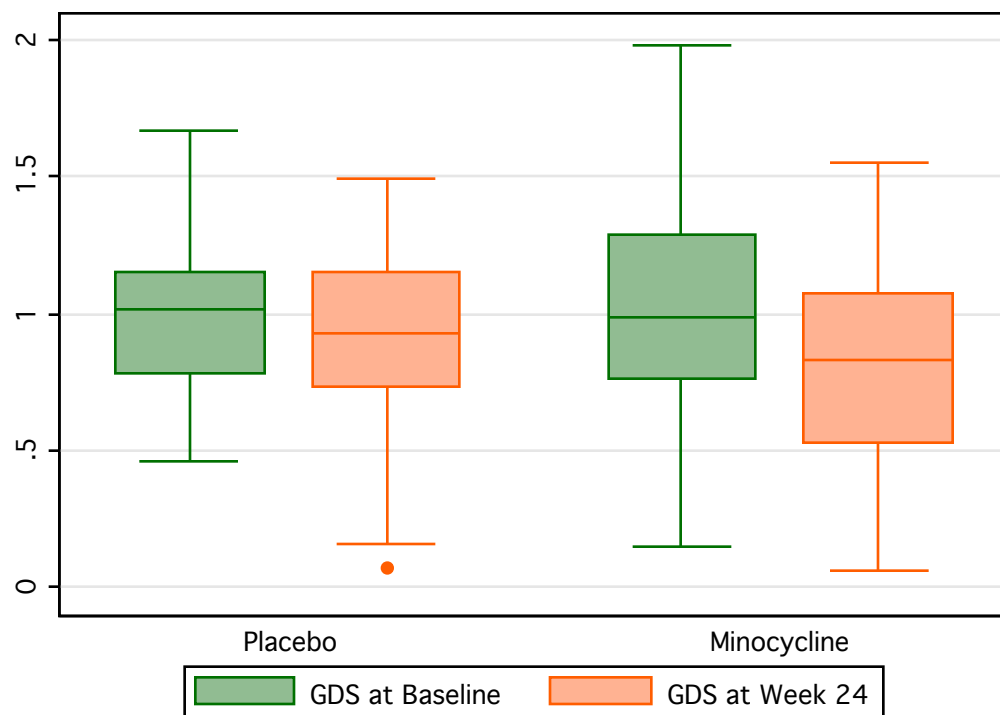
1. A simple t-test (this analysis scores mean GDS at week 24 between the two groups)
2. A change score t-test (this analysis calculates the change score from baseline to week 24 and then compares those scores between the two groups)
3. A linear regression model for week 24 GDS scores with predictors for treatment and baseline GDS
4. A GEE model which treats GDS values as repeated measures data and models how GDS scores change with time, treatment and a time-by-treatment interaction.

The data was simulated in a way which ensures that all four of these should all estimate the same quantity -- β_1 .

Q1: Among the 4 analysis, which would seem to be the most different in spirit from the other 3? Which, for you, seems to be simplest one to explain?

To me, #1 is the most different in spirit since it discards information about the baseline. Probably #1 is the easiest to explain and the one which relies most on the design of this as a clinical trial. Use of #2 is also very simple but I have found resistance to the use of "change scores" among colleagues. This surprises because, even though analysis #1 doesn't employ a change score and #2 does, the two analyses attempt to estimate the same quantity.

```
table rx, c(mean gds1 mean gds2)
-----
      rx | mean(gds1)  mean(gds2)
-----+-----
  Placebo |    .9847581    .909831
Minocycline |    1.035027    .8269244
-----
```



You can get a sense of the data using boxplots

`graph box gds1 gds2, over(rx) scheme(s1color)`

Analysis 1: t-test

```
. ttest gds2, by(rx)
```

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
Placebo	50	.909831	.0457161	.3232615	.8179611	1.001701
Minocyc	50	.8269244	.0531323	.3757024	.720151	.9336979
combined	100	.8683777	.035117	.3511698	.798698	.9380574
diff		.0829066	.0700928		-.0561904	.2220035
diff = mean(Placebo) - mean(Minocyc)				t =	1.1828	
Ho: diff = 0				degrees of freedom =	98	

Q2: What is the mean difference given by the t-test, can you identify it as a difference between two cells in the table of means? What is the two-sided p-value for the t-test?

The mean difference given by the t-test is -0.0829. I prefer to use the placebo group as a reference in the above which means subtracting mean of minocycline minus placebo. Thus, I change the sign from the above. This is the difference in the left column of the table of means 0.8269 - 0.9098.

Analysis 2: t-test on change score

Q3: What is the mean change in GDS from baseline to week 24 in each group?

```
table rx, c(mean diff)
```

rx	mean(diff)
Placebo	-.0749272
Minocycline	-.2081024

It is -0.07 in the placebo group and -0.21 in the minocycline group. One thing we learn is that GDS scores are declining in both groups.

Q4: What is the mean difference given by the t-test on differences? How does it differ from the one you got from the t-test under analysis 1? Why do they differ? What is the two-sided p-value for the test?

The difference is -0.13. It is difference of the mean changes on the two arms. It is also the difference of differences from baseline to 3 months in the two arms: $\text{diff placebo} = (0.909 - 0.985) - (1.035 - 0.827) = -0.1$. Conceptually, it is taking into account that the minocycline group started with higher GDS scores at baseline. That is why it has a higher average change than the ordinary t-test. However, we also see the SE shrink because the use of the baseline has amplified the precision of the treatment effect.

```
. ttest diff, by(rx)
```

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
Placebo	50	-.0749272	.0351813	.2487692	-.1456266	-.0042277
Minocyc	50	-.2081024	.0410878	.2905344	-.2906714	-.1255334
combined	100	-.1415148	.0277287	.2772869	-.1965345	-.086495
diff		.1331752	.0540918		.0258317	.2405187
diff = mean(Placebo) - mean(Minocyc)				t =	2.4620	
Ho: diff = 0				degrees of freedom =	98	
Ha: diff < 0		Ha: diff != 0		Ha: diff > 0		
Pr(T < t) = 0.9922		Pr(T > t) = 0.0156		Pr(T > t) = 0.0078		

Analysis 3: Linear Regression Model

Q5: What are the treatment effects given by model with and without gds1 as a predictor? Has the coefficient changed? Has the model R^2 changed?

The coefficient changes a little but the R^2 changes substantially because the baseline GDS is so predictive of GDS at follow-up. This reduces the SE of the rx coefficient. This analysis emulates the t-test on differences but does it in a different way so it gets a different coefficient. However, it gets a coefficient which is farther from 0 than the one which doesn't adjust for baseline because it is also taking into account the extremely slight difference in GDS scores at baseline.

```
. reg gds2 rx gds1
```

Source	SS	df	MS	Number of obs = 100		
Model	6.1446902	2	3.0723451	F(2, 97) = 49.15		
Residual	6.06401199	97	.062515588	Prob > F = 0.0000		
Total	12.2087022	99	.123320224	R-squared = 0.5033		
				Adj R-squared = 0.4931		
				Root MSE = .25003		

gds2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
rx	-.1180588	.0501354	-2.35	0.021	-.2175637	-.018554
gds1	.6992882	.0715417	9.77	0.000	.5572977	.8412787
_cons	.2212012	.078827	2.81	0.006	.0647514	.377651

Analysis 4: GEE Model

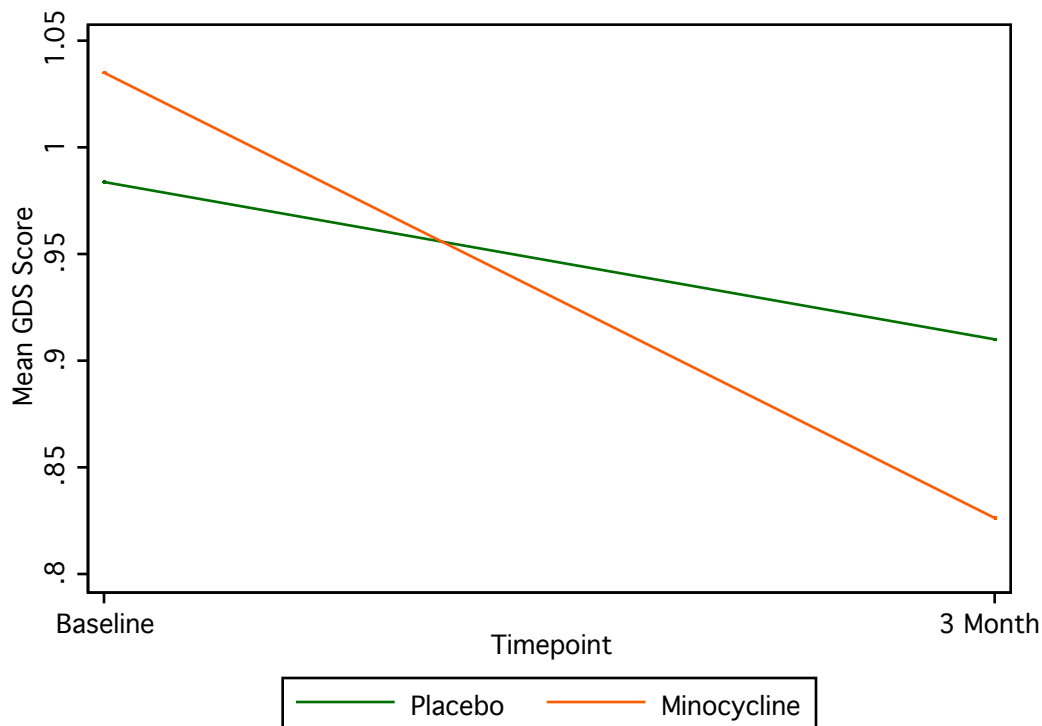
Q6: How do you interpret the results? How do they compare with the other models?

Perhaps this seemed complex since it involved repeated measures and an interaction. It is actually a very simple model which allows for the effect of treatment to vary by time. It thus defines four categories: baseline placebo, baseline minocycline, 3 month placebo and 3 month minocycline. The interaction term is simply the difference of differences of this mean and it EXACTLY the number we got for the t-test on the mean change scores. This will always be the case in the absence of missing values. You can see the fit of this model graphically by typing

predict out

```
twoway (line out timepoint if rx==0, sort) (line out timepoint if rx==1, sort), ytitle(Mean GDS Score) xtitle(Timepoint) xlabel(0 "Baseline" 1 "3 Months") legend(order(1 "Placebo" 2 "Minocycline")) scheme(s1color)
```

which gives



which shows the data in a very clear format.

4. SIMULATION

Q7: *Are all the method unbiased for estimating the coefficient beta ($=-0.15$)? What are the relative efficiency compared with the GE model? Do any of the methods appear to given an identical estimate of the treatment effect?*

They are all unbiased and so are all valid approaches and the point estimates are always identical for the ttest on differences and the GEE model. Then, it is natural to consider the issue of efficiency.

	rho = 0.70	rho = 0.50	rho = 0.00
t-test	1.69	1.02	0.50
t-test (diff)	1.00	1.00	1.00
regression	0.86	0.76	0.50

We see that the regression method is the most efficient in each setting, that the GEE and t-test do the same. We see that the t-test is a poor choice when the correlation between the baseline and 3 months is high, that it is equivalent to the difference and GEE when the correlation is 0.50 and it is a better choice when there is a less than 0.50 correlation. However, the regression performs well in all settings. The regression approach looks a little better here because the effect of baseline is truly linear (by the way I simulated the data) and hence there is no chance of a lack of model fit. Also, the regression method would fail completely if this was not a clinical trial. The best analysis is likely the GEE analysis. It has additional flexibility in handling the correlation. A still better choice would be a mixed model which will handle missing values better than some of the other methods.