

Annotated SAS Output

Program: Caffeine.sas
STAT 705, ANOVA Part 3.

Here is the program:

```
OPTIONS LS=80; /* restricts line size to 80 characters */
/* We will use @@ on the input statement to tell
   SAS to continue reading data on the same line */
DATA caffeine;
INPUT dose taps @@;
DATALINES;
    0 242    0 245    0 244    0 248    0 247
    0 248    0 242    0 244    0 246    0 242
100 248 100 246 100 245 100 247 100 248
100 250 100 247 100 246 100 243 100 244
200 246 200 248 200 250 200 252 200 248
200 250 200 246 200 248 200 245 200 250
;

PROC PRINT DATA=caffeine; RUN;

TITLE 'GLM';
PROC GLM DATA=caffeine PLOTS=diagnostics;
  CLASS dose (REF='0');
  MODEL taps = dose / SOLUTION ;
  LSMEANS dose / STDERR CL;
  MEANS dose / HOVTEST=BF; /* Brown-Forsythe test */
  RUN;

QUIT;
```

Obs	dose	taps
1	0	242
2	0	245
3	0	244
4	0	248
5	0	247
6	0	248
7	0	242
8	0	244
9	0	246
10	0	242
11	100	248
12	100	246
13	100	245
14	100	247
15	100	248
16	100	250
17	100	247
18	100	246
19	100	243
20	100	244
21	200	246
22	200	248
23	200	250
24	200	252
25	200	248
26	200	250
27	200	246
28	200	248
29	200	245
30	200	250

GLM

The GLM Procedure

Class Level Information		
Class	Levels	Values
dose	3	100 200 0

Number of Observations Read	30
Number of Observations Used	30

The first table generated by PROC GLM tells us the number of treatments and their values.

Pay attention to the order of the values – the reference level is always listed last. In our code, we have used the option REF='0', so the reference level is 0 (and is listed last).

In general, we do NOT change the reference level unless we have a very good reason for doing so.

It is always wise to check the number of observations. If the data has been read incorrectly by SAS, this could be your first indication that something is wrong.

For this experiment, we should have 10 observations for each dose, for a total of 30 observations. So it appears we are using the correct data.

This is the p-value for testing whether or not all the treatment means are equal. I call this the overall ANOVA F test.

The null hypothesis is

- For the cell means model . . . $H_0: \mu_1 = \mu_2 = \mu_3$
- For the effects model $H_0: \tau_1 = \tau_2 = 0$

Regardless of which model we are using, we reject H_0 (because $p=0.0062$, which is less than 0.05) and conclude at least one treatment has a different mean.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	61.4000000	30.7000000	6.18	0.0062
Error	27	134.1000000	4.9666667		
Corrected Total	29	195.5000000			

The MSE is 4.97.
This is an estimate for σ^2 .

R-Square	Coeff Var	Root MSE	taps Mean
0.314066	0.904098	2.228602	246.5000

Root MSE is 2.23.
This is the square root of MSE. It is an estimate for σ .

Ignore the Type I SS table

Source	DF	Type I SS	Mean Square	F Value	Pr > F
dose	2	61.40000000	30.70000000	6.18	0.0062

The Type III SS table will be important later

Source	DF	Type III SS	Mean Square	F Value	Pr > F
dose	2	61.40000000	30.70000000	6.18	0.0062

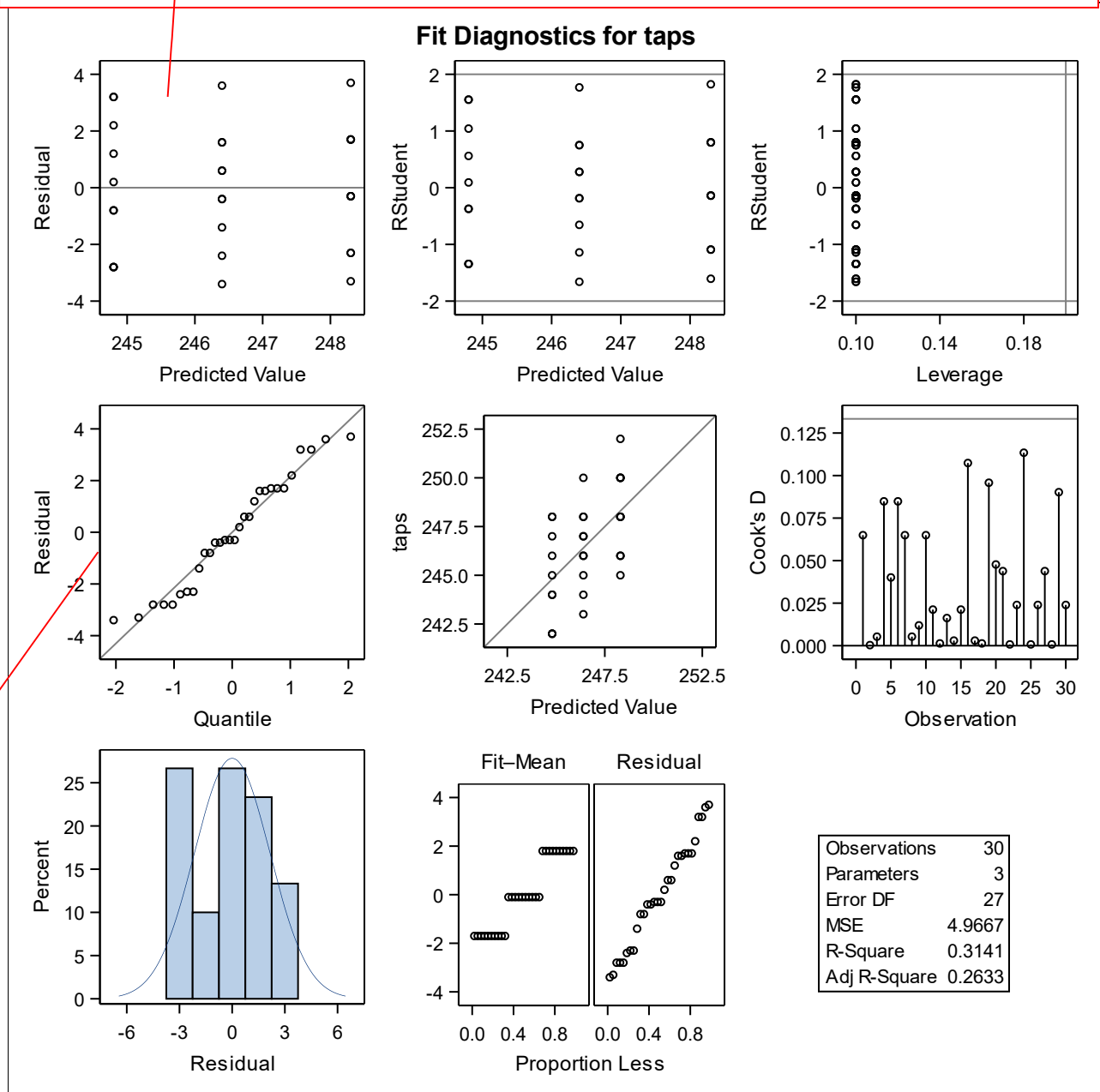
Parameter	Estimate		Standard Error	t Value	Pr > t
Intercept	244.8000000	B	0.70474582	347.36	<.0001
dose 100	1.6000000	B	0.99666109	1.61	0.1200
dose 200	3.5000000	B	0.99666109	3.51	0.0016
dose 0	0.0000000	B	.	.	.

This table gives the results of three tests for the effects model: $Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$

Intercept	$H_0: \mu = 0$ vs. $H_a: \mu \neq 0$	μ is the mean for the reference level, so this is testing whether or not the mean number of finger taps for dose 0 is equal to 0. We reject H_0 ($p<.0001$) and conclude it is not 0.
Dose 100	$H_0: \tau_1 = 0$ vs. $H_a: \tau_1 \neq 0$	τ_1 is the <u>effect</u> for dose 100 (as compared to the reference level), so this testing if the <u>difference</u> between dose 100 and dose 0 is equal to 0. We do not reject H_0 (because $p=0.12$, which is greater than 0.05). We conclude that there is no effect for dose 100. In other words, there is not a significant difference between the mean number of finger taps for dose 100 and the mean number of finger taps for dose 0.
Dose 200	$H_0: \tau_2 = 0$ vs. $H_a: \tau_2 \neq 0$	τ_2 is the <u>effect</u> for dose 200 (as compared to the reference level), so this testing if the <u>difference</u> between dose 200 and dose 0 is equal to 0. We reject H_0 ($p=0.0016$) and conclude the mean number of taps for dose 200 is not equal to the mean number of taps for dose 0.

GLM

This is the residual plot. We want the stacked points to have roughly the same vertical spread. If they do not, we will suspect that the assumption of equal variances has been violated. Use the results of the Brown-Forsythe test to make the final decision about this.



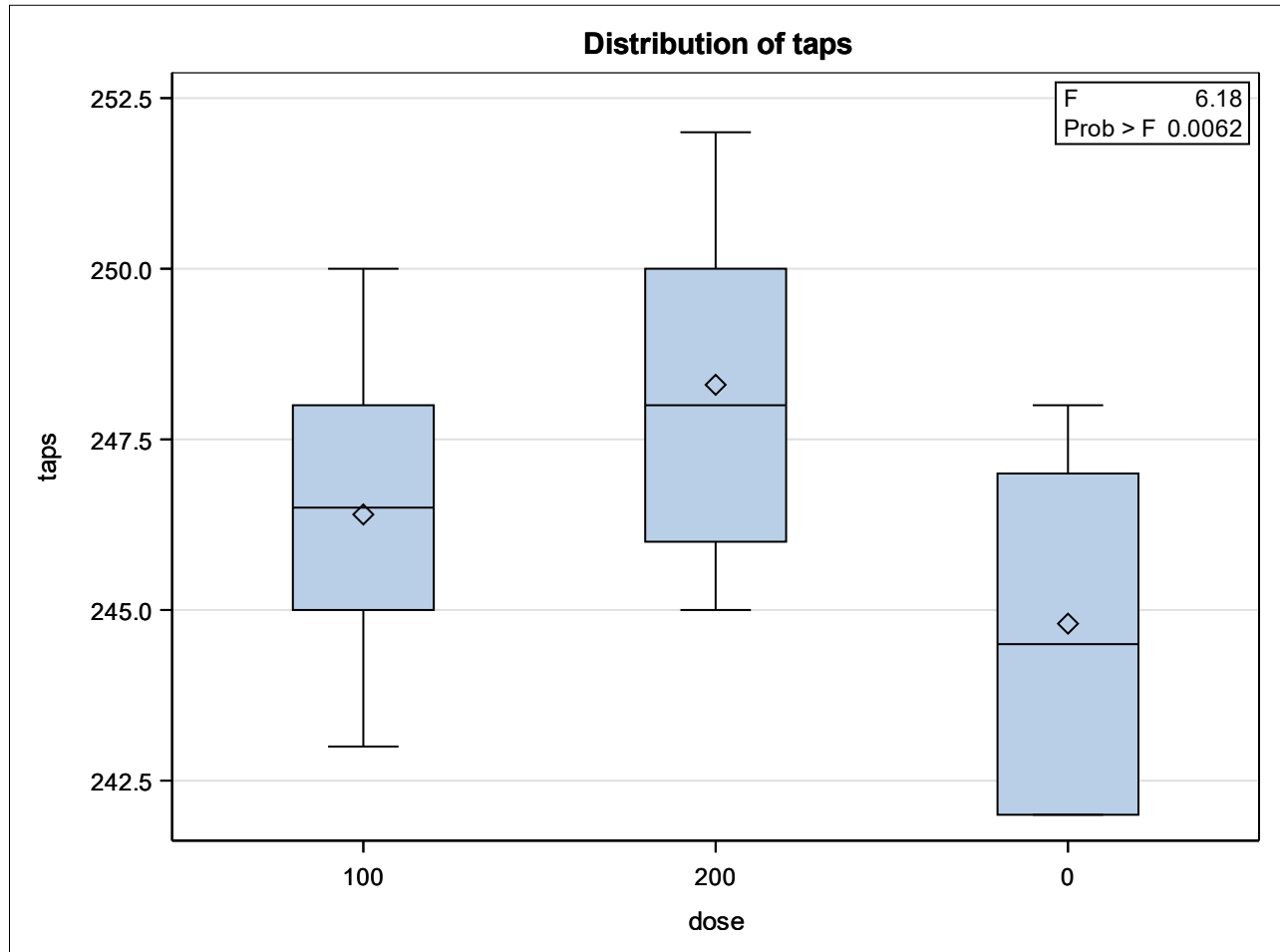
This is the normal probability plot. We want the points to follow the line. If they do not, we suspect the assumption of normality has been violated.

We can ignore (for now) the other plots on this page.

GLM

The GLM Procedure

Dependent Variable: taps



This is a nice plot, but it doesn't tell us much. Note that the reference level (0) is last, so the values for dose are not in numeric order. This will be important when we talk about contrasts later in the course.

**The GLS
Least Squares**

These are the estimates for the means and standard errors for the cell means model. We calculated these “by hand” in the last lesson.

dose	taps LSMEAN	Standard Error	Pr > t
100	246.400000	0.704746	<.0001
200	248.300000	0.704746	<.0001
0	244.800000	0.704746	<.0001

Hypotheses being tested

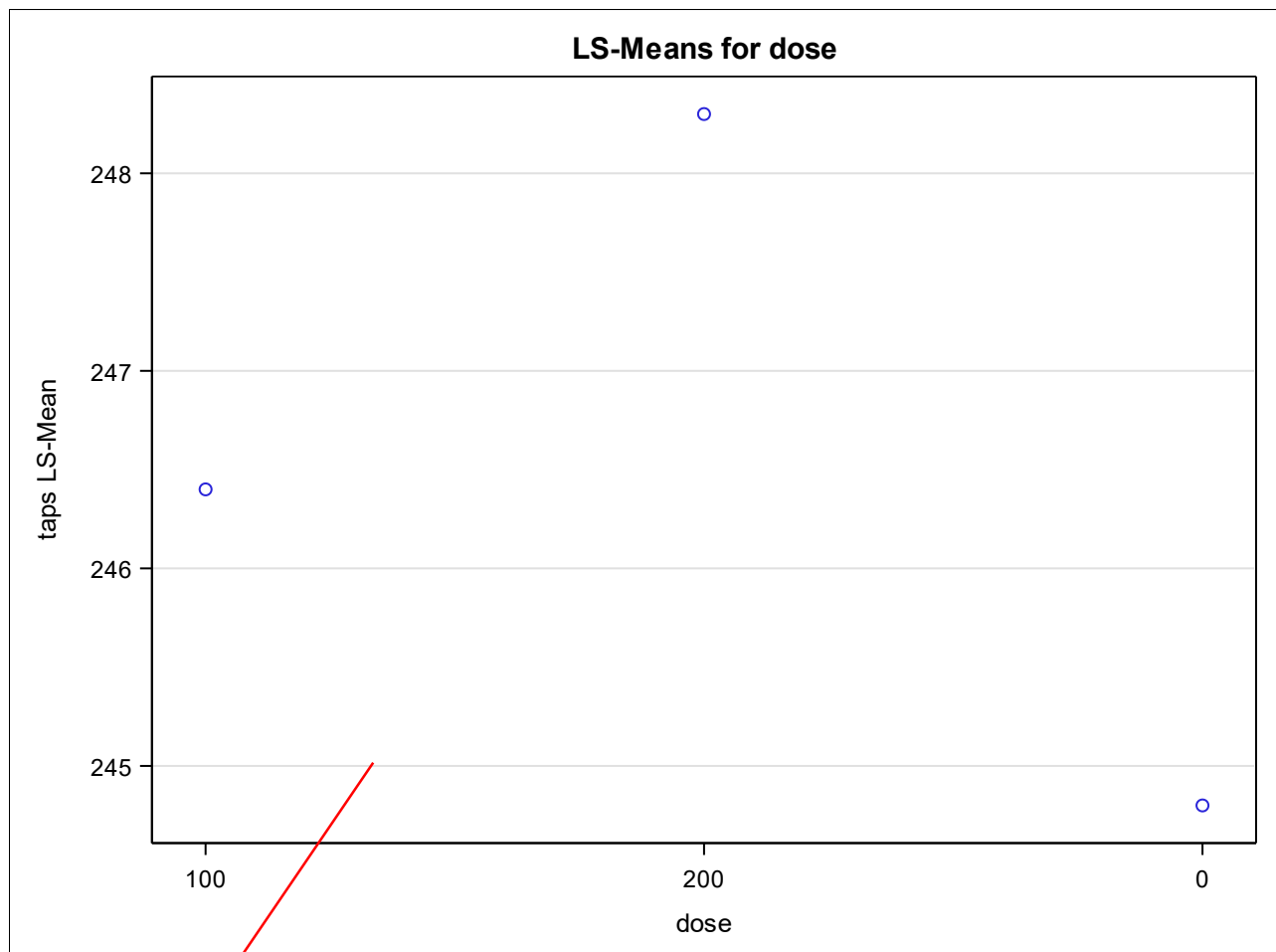
H0: $\mu_{100} = 0$ vs. Ha: $\mu_{100} \neq 0$

H0: $\mu_{200} = 0$ vs. Ha: $\mu_{200} \neq 0$

H0: $\mu_0 = 0$ vs. Ha: $\mu_0 \neq 0$

dose	taps LSMEAN	95% Confidence Limits	
100	246.400000	244.953981	247.846019
200	248.300000	246.853981	249.746019
0	244.800000	243.353981	246.246019

These are the 95% confidence intervals for the means that we calculated “by hand” in the last lesson.



This is a very uninformative graph. You can ignore this.

GLM

The GLM Procedure

Brown and Forsythe's Test for Homogeneity of taps Variance ANOVA of Absolute Deviations from Group Medians					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
dose	2	0.8667	0.4333	0.28	0.7565
Error	27	41.5000	1.5370		

Pay attention to this.

The Brown-Forsythe test is telling us whether or not we can believe that all treatment variances are equal.

H0 is all treatments have the same variance.

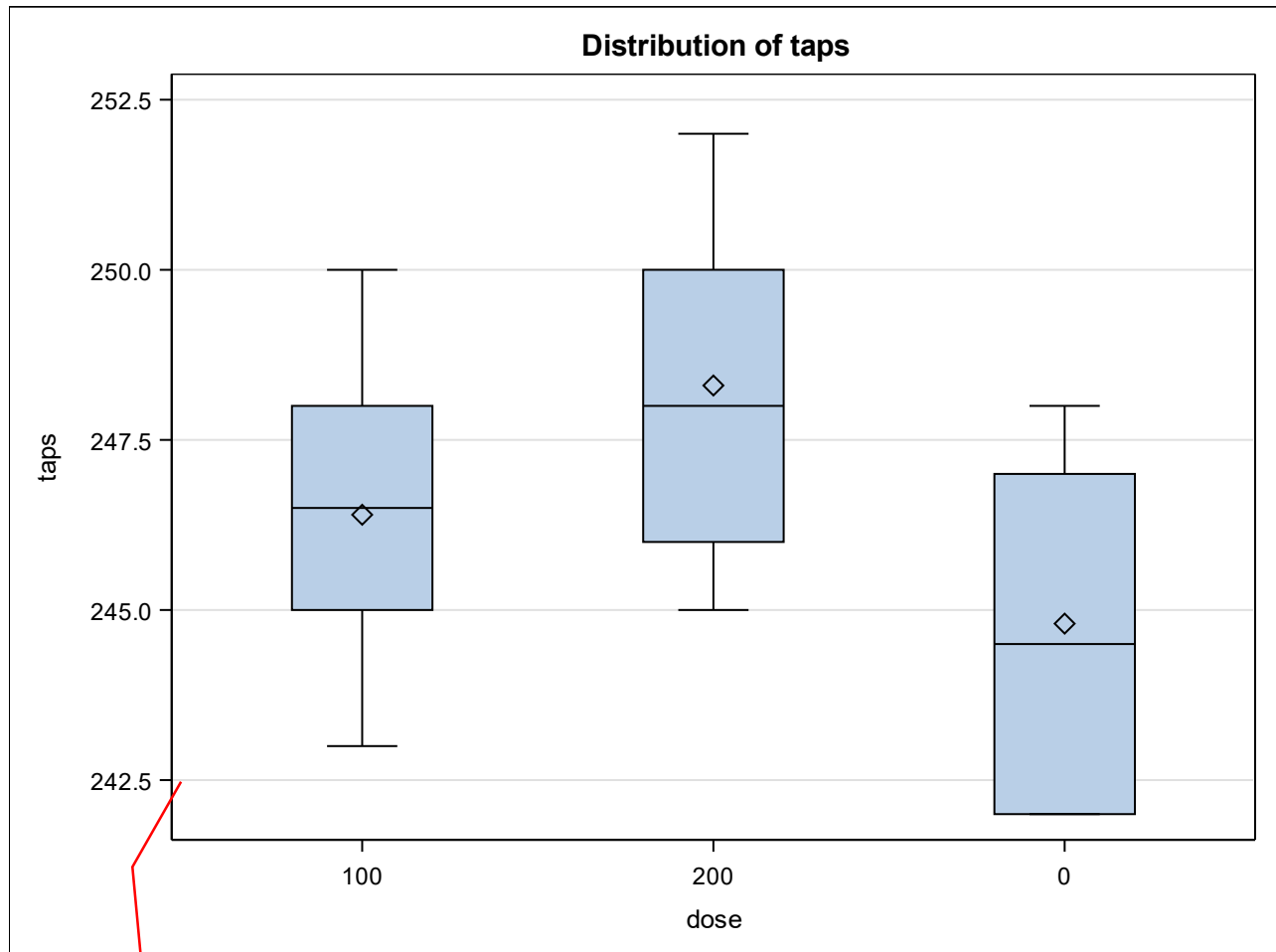
Ha is at least one treatment has a different variance.

Since $p = 0.7565$ is greater than 0.05, we do not reject H0.

We conclude the treatments all have the same variance, so the assumption of equal variances does not appear to be violated.

GLM

The GLM Procedure



This is a repeat of a previous graph that I said we could ignore. I don't know why SAS gives us two of these graphs. (sigh....)

Level of dose	N	taps	
		Mean	Std Dev
100	10	246.400000	2.06559112
200	10	248.300000	2.21359436
0	10	244.800000	2.39443800

This table is the result of the MEANS statement. We need the MEANS statement to get the Brown-Forsythe test, but we don't need to use this table. (The LSMEANS statement has already generated the table we will use.)