

Statistical Inference

1 Doing a Hypothesis Test for $B_1=0$ with PROC GLM

PROC GLM, as we have somewhat seen, is a powerful statement that can be used in sas. We have already used **PROC GLM** to obtain estimates for our slope parameter, and a 90, 95, and 99% confidence intervals. How do we obtain a test for $B_0=0$ or $B_1=0$? Well it turns out you already know how.

Here's an example, using the Hanes data set, that will perform the hypothesis test for $B_0=0$ and $B_1=0$, it will also obtain a 90% CI:

```
proc glm data=new;
by sex;
model sysbp=diabp\clparm alpha=.1;
output out=regdata r=resid p=yhat;
run;
quit;
```

```
proc gplot data=regdata;
plot yhat*resid;
run;
quit;
```

2 Power and sample size for test for a single slope

B_1^* = true slope parameter of interest. Have a sample of size n with B_1 = slope in the population. Now suppose we want to test the hypothesis.

$H_o : B_1^* = 0, H_A : B_1^* \neq 0$ Reject H_o if $T \leq R_1$ or $T \geq R_2$.

- R_1 is the negative critical value of t for a given alpha and df.
- R_2 is the positive critical value of t for a given alpha and df.

The test is based on that is has size α . The test statistic (T) is based on:

$$T = \frac{B_1 - 0}{SE_{B_1}} \quad (1)$$

We can use this test statistic to find the power of our test using the formula:

$$Power = ProbT_{df=n-1} \left(\frac{B_1}{SE_{B_1}} - R_2 \right) \quad (2)$$

If we are using R_2 or if we are using R_1 :

$$Power = 1 - ProbT_{df=n-1} \left(\frac{B_1}{SE_{B_1}} - R_1 \right) \quad (3)$$

Recall the formula for the standard error of B_1 :

$$SE_{B_1} = \frac{sd_y}{sd_x} \sqrt{\frac{1 - r_{xy}^2}{n - 2}} \quad (4)$$

Here is an example of finding power in SAS for a variety of values for B_1 and SE_{B_1} for a fix value of $n=42$:

```
data dummy;
do SEB1 = 3 to 5 by 1;
do B1 = -5 to 5 by 1;
R1 = -tinv(.975,40);
R2 =  tinv(.975,40);
power1 = 1-probt(B1/SEB1- R1,40);
power2 = probt(B1/SEB1- R2,40);
output;
end;
end;
run;

proc print;
run;
```

This could also be done for varying values of B_1 , sd_x , sd_y , r_{xy}^2 , and n instead of using SE_{B_1} .