Project Team 5

PAPER TOWEL EXPERIMENT

Fuxia Cheng,
Mehmet Kocak,
Yin Lau,
Hang Qui

STT 801
Design of Experiments
Fall 1999

Content

1. Problem Presentation

In this experiment, we tried to test the effects of **Brand** and **Printing** on the absorbency of paper towels. We have 3 different brands and 2 types of printing, namely, white and printed. Even though we consider the brand effect fixed, in fact, it can be random since we have more than 3 brands in the market. Under each treatment combination, we want to have 2 replicates. Running the experiment we want to measure two things: **Rate** at which the water droplets fell to the towel as our **covariant variable**, and **Absorbency:** # of water drops absorbed per square inch of towel as our **response variable**. Apparently, we will have 6 treatment combinations and under each of these, two replicates. Thus, our experiment structure will be like this:

		Printing	
		White	**Printed**
Brand	**1**	Experimental Unit-1 Experimental Unit-2	Experimental Unit-3 Experimental Unit-4
	2	Experimental Unit-5 Experimental Unit-6	Experimental Unit-7 Experimental Unit-8
	3	Experimental Unit-9 Experimental Unit-10	Experimental Unit-11 Experimental Unit-12

Then we want to set our models as follows:

Models

Model-1: Full Model for the 2 treatment factors (Brand and Printing), but no covariant-treatment interaction.

$$Y_{ijk} = m + t_{\mathrm{i}} + b_{j} + (t\ b)_{ij} + g\ (\mathrm{X}_{\mathrm{ijk}} - \overline{\mathrm{X}}_{...}) + e_{\mathrm{ijk}}$$

$$\text{where } e_{\mathrm{ijk}} \text{ iid N}(0,\ s^{2}),\ \ \mathrm{i} = 1, 2, 3;\ \ \mathrm{j} = 1, 2;\ \ \mathrm{k} = 1, 2$$

Model-2: includes only Factor A (Brand), ignores the levels of Factor B (Printing), but includes the Covariant-Treatment Interaction

$$\mathrm{Y}_{\mathrm{ik}} = m + t_{\mathrm{i}} + g_{\mathrm{i}}\,(\mathrm{X}_{\mathrm{ik}} - \overline{\mathrm{X}}_{\mathrm{i.}}) + \ + e_{\mathrm{ik}}$$

$$\text{where } e_{\mathrm{ik}} \text{ iid N}(0,\ s^{2}),\ \ \mathrm{i} = 1, 2, 3;\quad \mathrm{k} = 1, 2, 3, 4$$

2. Randomization

Before doing experiment, we want to decide the order of the runs. Therefore, we used the following random number to choose the order:

70997 79936 56865 05859 90**10**6 63553 40961 48235 **03**427 49626
69445 28663 72695 52180 **09**429 93969 52636 92737 88974 33488
3632**0 1**7617 30015 10365 61129 87529 85689 48237 52267 67689
93394 015**11** **07**119 97336 71**04**8 08178 77233 13976 47564 81056
97735 51085 **12**765 51821 51259 77452 16308 60756 92144 49442
02368 21382 52404 60268 89368 19885 55322 44819 01188 01011
54092 33362 94904 31273 04146 18594 29852 71585 52162 53916
46369 58586 23216 14513 83149 98736 23495 07056 97628 33787
09998 42698 **06**691 76988 13602 51851 48663 91245 85828 14346
09172 30168 90229 04734 59193 54164 58492 22421 74103 47070
25306 76468 26384 58151 32639 32363 **05**597

Thus, out randomized order is

10, 03, 09, 01, 11, 07, 04, 12, 02, 06, 05, 08

When we run the experiment, we will have a data table like following, where ****k**** means this experimental unit will be run in the k^{th} order:

		Printing	
		White	**Printed**
Brand	**1**	(X_{111}, Y_{111}) **4** (X_{112}, Y_{112}) **9**	X_{121}, Y_{121}) **2** (X_{122}, Y_{122}) **7**
	2	X_{211}, y_{211}) **11** (X_{212}, Y_{212}) **10**	X_{221}, Y_{221}) **6** (X_{222}, Y_{222}) **12**
	3	X_{311}, Y_{311}) **3** (X_{312}, Y_{312}) **1**	X_{321}, Y_{321}) **5** (X_{322}, Y_{322}) **8**

3. SAS Work

Data Set

First step is to assign a library named as “stt” under the a: direction. Then we created a permanent data set: stt.towel and input the data into this dataset. The variable “order” is the randomized order, “A” is factor A (Brand), “B” is factor B (Printing), “rate” is a covariate variable and “y” is the response variable: Absorbency. “x” is a recenterd covariate, which is equal to the rate minus the mean of the rate. After creating the data set stt.towel correctly, we analyzed both models using this dataset.

```
title1 "Paper Towels";
libname stt 'a:/';
data  towel;
  input order run A B rate Y;
  x=rate-2.026;
cards;
10  1 1 2 1.78  .7355
 3  2 2 2 1.867 .2828
 9  3 1 2 2.136 .3884
 1  4 1 1 1.952 .6777
11  5 3 1 1.8   .4364
 7  6 1 1 2.0   .6116
 4  7 2 2 2.071 .2929
12  8 3 2 1.951 .6465
 2  9 2 1 2.272 .2525
 6 10 2 1 2.25  .2727
 5 11 3 2 2.075 .6707
 8 12 3 1 2.158 .3313
;
run;

proc print;
run;
```

Model 1

In model 1 we used “proc glm”. Factor A and factor B, AB interaction and covariate are included in the model statement. Due to the significant interaction between main effects, we could not conclude the effects of A and B via the output. The effect of A and B may be canceled out or be reduced. We used “lsmeans” statement and “slice =” option to test the specified effects: B effect within A and A effect within B. We are mainly interested in the factor A, so we used “lsmeans” statement and “pdiff=all” option to do the multiple comparison in factor A. “pdiff” requests the p-values for differences of the least squares means. “all” requests all pairwise differences. The “cl” option requests that confidence limits be constructed for each of the least-squares means. The confidence level is 0.95 by default. Tukey is used as adjustment. The “contrast” statement provides a test for difference of means and “estimate” statement gives the estimated parameter of tau and slope of covariate. “proc plot” gives several graphs we need.

```
proc glm data=stt.towel;
 title2 'Model 1';
 class A B;
 model y=a b a*b x;
 output out=b predicted=ypred
residual=res;
run;

 lsmeans a*b/slice=a;
 lsmeans a*b/slice=b;
 lsmeans a /pdiff=all cl adjust=tukey;

 contrast 'a1-a2' a 1 -1  0;
 contrast 'a1-a3' a 1  0 -1;
 contrast 'a2-a3' a 0  1 -1;

 estimate 'tau1' a 2 -1 -1 /divisor=3;
 estimate 'tau2' a -1 2 -1/divisor=3;
 estimate 'slope' x 1;
run;

proc plot;
   plot res*ypred;
   plot ypred*a;
   plot y*x;
run;
```

Model 2

In model 2 we still use “proc glm” to analyze the effects but ignore the factor B and added the interaction between factor A and covariate. Thus the model statement contains factor A, covariate X and the interaction between A and X. We used option “solution” in the model statement to print a solution to parameter estimates. It gives us the individual slopes of covariate in different A levels. If the interaction of factor A and covariate is significant, these slopes are not equal. In other words, the lines are not parallel in different A levels. We also used “estimate” and “proc plot” statements to get the estimated parameters and graphs. We checked the model adequacy in model 2. “proc univariate plot normal” tests the normality assumption of residuals and the normal probability plot of the residuals.

```
proc sort data=stt.towel; by order;
run;

proc glm data=stt.towel;
  title2 'Model 2';
  class A;
  model y=A x A*x/solution;
output out=b predicted=ypred
residual=res;
run;

estimate "1-2" A 1 -1 0;
estimate "1-3" A 1 0 -1;
estimate '2-3' A 0 1 -1;

estimate 'tau2' a -1 2 -1/divisor=3;
estimate 'slope' x 1;
run;

proc univariate plot normal;
   var res;
run;

proc plot;
   plot res*order;
   plot res*ypred;
   plot y*x;
run;
```

Model 3

In Model 3, we tested the alternative design, which proposed a model same as Model 1, but allowing different slopes. We used “solution” to get ther F-test and the corresponding p-values.

```
proc sort data=stt.towel; by order;
run;

proc glm data=stt.towel;
 title2 'Model 3';
 class A;
 model y=A B A*B x A*x B*x/solution;
output out=b predicted=ypred
residual=res;
run;
```

4. Output

Model 1

Paper Towel Experiment
Model 1

The GLM Procedure

Class Level Information

Class	Levels	Values
A	3	1 2 3
B	2	1 2

Number of observations 12

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	0.35669402	0.05944900	10.64	0.0101
Error	5	0.02794574	0.00558915		
Corrected Total	11	0.38463976			

R-Square	Coeff Var	Root MSE	Y Mean
0.927346	16.02299	0.074761	0.466583

Source	DF	Type I SS	Mean Square	F Value	Pr > F
A	2	0.23318068	0.11659034	20.86	0.0037
B	1	0.01573976	0.01573976	2.82	0.1542
A*B	2	0.06722465	0.03361233	6.01	0.0467
X	1	0.04054892	0.04054892	7.25	0.0431

Source	DF	Type III SS	Mean Square	F Value	Pr > F
A	2	0.10475260	0.05237630	9.37	0.0204
B	1	0.00170997	0.00170997	0.31	0.6040
A*B	2	0.09321477	0.04660739	8.34	0.0255
X	1	0.04054892	0.04054892	7.25	0.0431

Paper Towel Experiment

Model 1

The GLM Procedure

Least Squares Means

A	B	Y LSMEAN
1	1	0.61926717
1	2	0.52742935
2	1	0.38189932
2	2	0.25891357
3	1	0.35999014
3	2	0.65200046

A*B Effect Sliced by A for Y

A	DF	Sum of Squares	Mean Square	F Value	Pr > F
1	1	0.008417	0.008417	1.51	0.2744
2	1	0.009810	0.009810	1.76	0.2426
3	1	0.084648	0.084648	15.15	0.0115

A*B Effect Sliced by B for Y

B	DF	Sum of Squares	Mean Square	F Value	Pr > F
1	2	0.076188	0.038094	6.82	0.0373
2	2	0.160182	0.080091	14.33	0.0085

Paper Towel Experiment
Model 1

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

A	Y LSMEAN	LSMEAN Number
1	0.57334826	1
2	0.32040644	2
3	0.50599530	3

Least Squares Means for effect A
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: Y

i/j	1	2	3
1		0.0187	0.4696
2	0.0187		0.0512
3	0.4696	0.0512	

A	Y LSMEAN	95% Confidence Interval	
1	0.573348	0.473097	0.673599
2	0.320406	0.215086	0.425727
3	0.505995	0.408813	0.603177

Least Squares Means for Effect A

i	j	Difference Between Means	Simultaneous 95% Confidence Interval for LSMean(i)-LSMean(j)	
1	2	0.252942	0.058453	0.447431
1	3	0.067353	-0.105575	0.240281
2	3	-0.185589	-0.372441	0.001264

```
                         Paper Towel Experiment
                                Model 1
Dependent Variable: Y

      Contrast        DF     Contrast SS     Mean Square    F Value    Pr > F
      a1-a2            1      0.10009092      0.10009092      17.91    0.0082
      a1-a3            1      0.00897688      0.00897688       1.61    0.2609
      a2-a3            1      0.05837809      0.05837809      10.44    0.0232

Dependent Variable: Y
                                        Standard
      Parameter        Estimate         Error           t Value    Pr > |t|

      tau1              0.10676492      0.03248352        3.29      0.0218
      tau2             -0.14617689      0.03482672       -4.20      0.0085
      slope            -0.50765667      0.18847476       -2.69      0.0431
```

In Model 1, we test the effect of A(Brand), B(Printing), A*B(interaction of Brand and Printing), and X(covariate -- rate of absorbency). Since both A and A*B have significant effects, we did the lsmeans, multiple comparison, and contrast estimate to further analyze the A effect. Following is a summary of the output for model one:

ANOVA Table

Table 1 -- Result from proc glm

	SS	d.f.	MS	F	p
A	**0.10475**	**2**	**0.5237**	**9.37**	**0.0204**
B	0.00171	1	0.00171	0.31	0.6040
A*B	**0.04661**	**2**	**0.04661**	**8.34**	**0.0255**
X	**0.04054**	**1**	**0.04055**	**7.25**	**0.0431**
Error	0.02795	5	0.00559		

Table 2 -- Result from lsmeans

	SS	d.f.	MS	F	p
A(B1)	**0.07619**	**2**	**0.03809**	**6.82**	**0.0373**
A(B2)	**0.16018**	**2**	**0.08009**	**14.33**	**0.0085**
B(A1)	0.00842	1	0.08417	1.51	0.2744
B(A2)	0.00981	1	0.00098	1.76	0.2426
B(A3)	**0.084648**	**1**	**0.08465**	**15.15**	**0.0115**

Note: B within A: Only significant with A3

A within B: Significant at both levels of B

Output for Multiple Comparisons

This part of output is obtained by lsmeans a /pdiff=all cl adjust=tukey;

A	Y LSMEAN	95% Confidence Interval	
1	0.57334826	0.473097	0.673599
2	0.32040644	0.215086	0.425727
3	0.50599530	0.408813	0.603177

H_0: LSMean(A_i)= LSMean(A_j)		p-value	Difference between means	Simultaneous 95% Confidence Interval	
A_1	**A_2**	**0.0187**	**0.252942**	**0.058453**	**0.447431**
A_1	A_3	0.4696	0.067353	-0.105575	0.240281
A_2	**A_3**	**0.0512**	**-0.185589**	**-0.372441**	**0.001264**

Contrast Estimates

	SS	d.f.	MS	F	p
A1-A2	**0.10009**	**1**	**0.10009**	**17.91**	**0.0082**
A1-A3	0.00898	1	0.00898	1.61	0.2609
A2-A3	**0.05838**	**1**	**0.05838**	**10.44**	**0.0232**

Parameter	Estimate	p
τ_1	0.10676	0.0082
τ_2	-0.14618	0.2609
γ	-0.50766	0.0232

Note: A2 is the most significant one

Graphs

1. Res*Ypred – There is a curve pattern in the graph. It indicates that there is probably a significant effect in the interactions.

```
                 Plot of res*ypred.  Legend: A = 1 obs, B = 2 obs, etc.

     res ‚
   0.100 ˆ
         ‚
         ‚
         ‚                                                                  A
         ‚
   0.075 ˆ
         ‚
         ‚
         ‚
         ‚       A
   0.050 ˆ
         ‚                                                               A
         ‚               A
         ‚
         ‚
   0.025 ˆ
         ‚                                                                   A
         ‚
         ‚
         ‚            A
   0.000 ˆ
         ‚          A
         ‚
         ‚
         ‚                                                                A
  -0.025 ˆ
         ‚
         ‚
         ‚                                         A
         ‚                                                                        A
  -0.050 ˆ
         ‚                      A
         ‚
         ‚
         ‚
  -0.075 ˆ
         ‚
         ‚                                         A
         ‚
         ‚
  -0.100 ˆ
         ‚
         Šƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒ
           0.2           0.3           0.4           0.5           0.6           0.7
                                             ypred
```

2. Ypred*A – Different levels of A(Brands) have different variances. Particularly, A2 is significantly different from A1 and A3. But A1, and A3 also don't have similar variance.

```
              Plot of ypred*A.  Legend: A = 1 obs, B = 2 obs, etc.

ypred ‚
      ‚
  0.7 ˆ
      ‚                                                                                          A
      ‚
      ‚B
      ‚
      ‚A                                                                                         A
      ‚
  0.6 ˆ
      ‚
      ‚
      ‚
      ‚
      ‚
      ‚
  0.5 ˆ
      ‚
      ‚A                                                                                         A
      ‚
      ‚
      ‚
      ‚
  0.4 ˆ
      ‚
      ‚
      ‚
      ‚                                                 A
      ‚
      ‚
  0.3 ˆ                                                                                          A
      ‚
      ‚                                                 A
      ‚                                                 A
      ‚                                                 A
      ‚
      ‚
  0.2 ˆ
      ‚
      Šˆƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒˆ
       1                                            2                                            3

                                                        A
```

3. Y*X – It looks like follows the normal pattern pretty closely. Generally, the higher X (covariate, rate) is, the Lower Y (absorbency) becomes. Additional label according to the treatments are added to further analysis if there is any significant relations between the factors. Based on the graph, we can see that for A1 (Brand 1), the result doesn't seems to vary because of B(Printing). It only follows the general normal pattern we just mentioned. For A2(Brand 2), the results are low (and around the same level) no matter how B(Printing) and X(Rate) change. For A3, the graph reflects that higher results can be obtained with B2(Printed) instead of B1(White). And this goes with our result from LSMEANS that B is only significant within A3.

```
                  Plot of Y*X.  Legend: A = 1 obs, B = 2 obs, etc.

  Y ,
    ,
0.8 ˆ
    ,
    ,
    ,       A(1,2)
    ,
0.7 ˆ
    ,                           A(1,1)       A(3,2)
    ,
    ,                           A(3,2)
    ,                                  A(1,1)
0.6 ˆ
    ,
    ,
    ,
    ,
0.5 ˆ
    ,
    ,
    ,          A(3,1)
    ,
0.4 ˆ
    ,                                                 A(1,2)
    ,
    ,                                                    A(3,1)
    ,
0.3 ˆ                                          A(2,2)
    ,                 A(2,2)                                        A(2,1)
    ,                                                                  A(2,1)
    ,
    ,
0.2 ˆ
    ,
    Šfˆffffffffffffˆffffffffffffˆffffffffffffˆffffffffffffˆffffffffffffˆffffffffffffˆff
    -0.3        -0.2        -0.1         0.0         0.1         0.2         0.3

                                         X
```

Model 2

Paper Towel Experiment
Model 2

The GLM Procedure
Class Level Information

Class	Levels	Values
A	3	1 2 3

Number of observations 12

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	0.29351187	0.05870237	3.87	0.0652
Error	6	0.09112789	0.01518798		
Corrected Total	11	0.38463976			

R-Square	Coeff Var	Root MSE	Y Mean
0.763082	26.41319	0.123240	0.466583

Source	DF	Type I SS	Mean Square	F Value	Pr > F
A	2	0.23318068	0.11659034	7.68	0.0222
X	1	0.02392475	0.02392475	1.58	0.2561
A*X	2	0.03640643	0.01820322	1.20	0.3648

Source	DF	Type III SS	Mean Square	F Value	Pr > F
A	2	0.13656672	0.06828336	4.50	0.0641
X	1	0.03286056	0.03286056	2.16	0.1917
A*X	2	0.03640643	0.01820322	1.20	0.3648

Parameter		Estimate		Standard Error	t Value	Pr > |t|
Intercept		0.5179459226	B	0.06312238	8.21	0.0002
A	1	0.0290494941	B	0.09271872	0.31	0.7647
A	2	-.2372250801	B	0.09440880	-2.51	0.0457
A	3	0.0000000000	B	.	.	.
X		-.1093025807	B	0.45636133	-0.24	0.8187
A*X	1	-.8450123906	B	0.66519851	-1.27	0.2510
A*X	2	0.0475515413	B	0.59255829	0.08	0.9386
A*X	3	0.0000000000	B	.	.	.

NOTE: The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.

Paper Towel Experiment
Model 2

Dependent Variable: Y

Parameter	Estimate	Standard Error	t Value	Pr > |t|
1-2	0.26627457	0.09767760	2.73	0.0344
1-3	0.02904949	0.09271872	0.31	0.7647
2-3	-0.23722508	0.09440880	-2.51	0.0457

Level of		--------------Y--------------		--------------X--------------	
A	N	Mean	Std Dev	Mean	Std Dev
1	4	0.60330000	0.15194637	-0.05900000	0.14701927
2	4	0.27522500	0.01724903	0.08900000	0.18824983
3	4	0.52122500	0.16462421	-0.03000000	0.15591237

The UNIVARIATE Procedure
Variable: res

Moments

N	12	Sum Weights	12
Mean	0	Sum Observations	0
Std Deviation	0.09101843	Variance	0.00828435
Skewness	-0.0844311	Kurtosis	0.26095646
Uncorrected SS	0.09112789	Corrected SS	0.09112789
Coeff Variation	.	Std Error Mean	0.02627476

Basic Statistical Measures

Location		Variability	
Mean	0.00000	Std Deviation	0.09102
Median	-0.00096	Variance	0.00828
Mode	.	Range	0.33033
		Interquartile Range	0.09988

Tests for Location: Mu0=0

Test	-Statistic-		-----p Value------	
Student's t	t	0	Pr > |t|	1.0000
Sign	M	0	Pr >= |M|	1.0000
Signed Rank	S	0	Pr >= |S|	1.0000

Tests for Normality

Test	--Statistic---		-----p Value------	
Shapiro-Wilk	W	0.984549	Pr < W	0.9958
Kolmogorov-Smirnov	D	0.111224	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.025378	Pr > W-Sq	>0.2500
Anderson-Darling	A-Sq	0.162803	Pr > A-Sq	>0.2500

Paper Towel Experiment
Model 2

Quantiles (Definition 5)

Quantile	Estimate
100% Max	0.158109904
99%	0.158109904
95%	0.158109904
90%	0.120356384
75% Q3	0.049938835
50% Median	-0.000963934
25% Q1	-0.049938835
10%	-0.106248306
5%	-0.172217982
1%	-0.172217982
0% Min	-0.172217982

Extreme Observations

-------Lowest------		------Highest------	
Value	Obs	Value	Obs
-0.1722180	8	0.0149580	4
-0.1062483	11	0.0397924	7
-0.0536208	9	0.0600853	1
-0.0462569	10	0.1203564	12
-0.0130301	2	0.1581099	5

```
 Stem Leaf                     #             Boxplot
  1 6                          1                |
  1 2                          1                |
  0 6                          1             +-----+
  0 114                        3             |  +  |
 -0 11                         2             *-----*
 -0 55                         2             +-----+
 -1 1                          1                |
 -1 7                          1                |
    ----+----+----+----+
                Multiply Stem.Leaf by 10**-1
```

```
                   Paper Towel Experiment
                          Model 2

Variable:  res

Normal Probability Plot
0.175+                                         *+++++
|                                    *+++++
|                               ++*++
|                         +*+*+*
|                    *+*+*
|               ++*++
|         +++++*
-0.175+    +++++*
+----+----+----+----+----+----+----+----+----+----+
-2        -1         0        +1        +2
```

In Model 2, when we take the covariate effect into account, we find that none of the test are significant, including A(Brand), but one things that should be noted is that if we look at the Type I result, A effect is still very significant. Here is a summary of the result:

ANOVA Table

	SS	d.f.	MS	F	p
A (Type I)	**0.23318**	**2**	**0.11659**	**7.68**	**0.0222**
A	0.13657	2	0.06828	4.50	0.0641
X	0.03286	1	0.03286	2.16	0.1917
A*X	0.03641	2	0.01820	1.20	0.3648
Error	0.09113	6	0.01519		

1. Note that none is significant looking at Type III, however, when we look at Type I, A effect is significant.
   Type I F= 7.68, p = 0.0222
   Type III F= 4.50, p = 0.0641

2. Since the covariate can affect the result, we tried to test on the different slopes by using the option solution in SAS.

3. The UNIVARIATE Procedure is run on the residual for the purpose of model adequacy checking, which will be explained in the next section.

Graphs

1. Res*Ypred – There is no significant pattern in the graph. No specific conclusion can be drawn from the graph

```
                      Plot of res*ypred.  Legend: A = 1 obs, B = 2 obs, etc.

  res ,
 0.20 ˆ
      ,
      ,
      ,
      ,                                                A
 0.15 ˆ
      ,
      ,
      ,                                                  A
      ,
 0.10 ˆ
      ,
      ,
      ,
      ,                                                               A
 0.05 ˆ
      ,                                                        A
      ,
      ,
      ,           A A
 0.00 ˆ
      ,           A   A
      ,
      ,
      ,
-0.05 ˆ                                      A                                                A
      ,
      ,
      ,
      ,
-0.10 ˆ
      ,                                                    A
      ,
      ,
      ,
-0.15 ˆ
      ,
      ,                                              A
      ,
      ,
-0.20 ˆ
      ,
      Šƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒ
        0.2           0.3           0.4           0.5           0.6           0.7           0.8

                                                   ypred
```

2. Ypred*A – Different levels of A(Brands) have different variances. Particularly, A2 is significantly different from A1 and A3. But A1, and A3 also don't have similar variance. This is expected as we already know that when only looking at A (Brand) by itself, ignoring other variations, it does have significant impact on Y (absorbency). But as we know that when the effect of the covariate is taking into account, the p-value reflects that it's not significant, there this graph doesn't really help much in drawing our conclusion.

```
                        Plot of Y*A.  Legend: A = 1 obs, B = 2 obs, etc.
0.8 ˆ
    ,
    ,
    ,
    ,
    ,A
    ,
0.7 ˆ
    ,
    ,A                                                                                     A
    ,
    ,                                                                                      A
    ,
    ,A
0.6 ˆ
    ,
    ,
  Y ,
    ,
    ,
    ,
0.5 ˆ
    ,
    ,
    ,
    ,                                                                                      A
    ,
    ,
0.4 ˆ
    ,A
    ,
    ,
    ,
    ,                                                                                      A
    ,
0.3 ˆ                                           A
    ,                                           A
    ,                                           A
    ,                                           A
    ,
    ,
    ,
0.2 ˆ
    Šˆffffffffffffffffffffffffffffffffffffffffffˆffffffffffffffffffffffffffffffffffffffffffˆ
     1                                          2                                          3
                                                A
```

3. Res*Order – There is no significant pattern in the graph, which means that the order of running the experiment probably won’t affect the output of the experiment.

```
              Plot of res*ORDER.  Legend: A = 1 obs, B = 2 obs, etc.

      res ,
          ,
     0.20 ˆ
          ,
          ,
          ,                             A
     0.15 ˆ
          ,
          ,                                                                              A
          ,
     0.10 ˆ
          ,
          ,
          , A
     0.05 ˆ
          ,                                           A
          ,
          ,                      A
     0.00 ˆ                                    A
          ,        A      A
          ,
          ,
    -0.05 ˆ                                                         A      A
          ,
          ,
          ,
    -0.10 ˆ                                                                       A
          ,
          ,
          ,
    -0.15 ˆ
          ,
          ,                                                  A
          ,
    -0.20 ˆ
          ,
          Šƒˆƒƒƒƒƒƒˆƒƒƒƒƒƒˆƒƒƒƒƒƒˆƒƒƒƒƒƒˆƒƒƒƒƒƒˆƒƒƒƒƒƒˆƒƒƒƒƒƒˆƒƒƒƒƒƒˆƒƒƒƒƒƒˆƒƒƒƒƒƒˆƒƒƒƒƒƒˆƒƒƒƒƒƒˆƒ
            1      2      3      4      5      6      7      8      9     10     11     12

                                               ORDER
```

4. Y*X – It looks like follows the normal pattern pretty closely. Generally, the higher X (covariate, rate) is, the Lower Y (absorbency) becomes.

```
                      Plot of Y*X.  Legend: A = 1 obs, B = 2 obs, etc.

   Y ,
     ,
 0.8 ˆ
     ,
     ,
     ,        A
     ,
 0.7 ˆ
     ,                              A               A
     ,
     ,                              A
     ,                                     A
 0.6 ˆ
     ,
     ,
     ,
     ,
 0.5 ˆ
     ,
     ,
     ,           A
     ,
 0.4 ˆ
     ,                                                      A
     ,
     ,                                                         A
     ,
 0.3 ˆ                                              A
     ,                   A                                                 A
     ,                                                                        A
     ,
     ,
 0.2 ˆ
     ,
     Šƒˆƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒƒˆƒƒ
    -0.3         -0.2         -0.1          0.0          0.1          0.2          0.3

                                              X
```

Model 3

Paper Towel Experiment
Model 3

The GLM Procedure
Class Level Information

Class	Levels	Values
A	3	1 2 3

Number of observations 12

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	9	0.38457642	0.04273071	1349.28	0.0007
Error	2	0.00006334	0.00003167		
Corrected Total	11	0.38463976			

R-Square	Coeff Var	Root MSE	Y Mean
0.999835	1.206118	0.005628	0.466583

Source	DF	Type I SS	Mean Square	F Value	Pr > F
A	2	0.23318068	0.11659034	3681.50	0.0003
B	1	0.01573976	0.01573976	497.00	0.0020
A*B	2	0.06722465	0.03361233	1061.35	0.0009
X	1	0.04054892	0.04054892	1280.39	0.0008
A*X	2	0.02589914	0.01294957	408.90	0.0024
B*X	1	0.00198326	0.00198326	62.62	0.0156

Source	DF	Type III SS	Mean Square	F Value	Pr > F
A	2	0.06204697	0.03102349	979.61	0.0010
B	1	0.00184310	0.00184310	58.20	0.0168
A*B	2	0.07232635	0.03616317	1141.90	0.0009
X	1	0.00421449	0.00421449	133.08	0.0074
A*X	2	0.02371772	0.01185886	374.46	0.0027
B*X	1	0.00198326	0.00198326	62.62	0.0156

Paper Towel Experiment
Model 3
The GLM Procedure

Parameter		Estimate		Standard Error	t Value	Pr > |t|
Intercept		0.078923297	B	0.00915835	8.62	0.0132
A	1	0.568202936	B	0.01360119	41.78	0.0006
A	2	0.365874473	B	0.03854837	9.49	0.0109
A	3	0.000000000	B	.	.	.
B		0.291118551	B	0.00576580	50.49	0.0004
B*A	1	-0.366802844	B	0.00826860	-44.36	0.0005
B*A	2	-0.368337678	B	0.02071697	-17.78	0.0031
B*A	3	0.000000000	B	.	.	.
X		-0.784534706	B	0.07179777	-10.93	0.0083

Dependent Variable: Y

Parameter		Estimate		Standard Error	t Value	Pr > |t|
X*A	1	-1.170370741	B	0.06226229	-18.80	0.0028
X*A	2	-0.152927167	B	0.07023657	-2.18	0.1614
X*A	3	0.000000000	B	.	.	.
B*X		0.490744241		0.06201323	7.91	0.0156

NOTE: The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.

Dependent Variable: Y

Parameter	Estimate	Standard Error	t Value	Pr > |t|
1-2	0.20232846	0.04239385	4.77	0.0412
1-3	0.56820294	0.01360119	41.78	0.0006
2-3	0.36587447	0.03854837	9.49	0.0109

Level of		------------Y----------		-----------B-----------		-----------X-----------	
A	N	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
1	4	0.60330000	0.15194637	1.50000000	0.57735027	-0.05900000	0.14701927
2	4	0.27522500	0.01724903	1.50000000	0.57735027	0.08900000	0.18824983
3	4	0.52122500	0.16462421	1.50000000	0.57735027	-0.03000000	0.15591237

Parameter	Estimate	Standard Error	t Value	Pr > |t|
slope	-1.22563401	0.10624478	-11.54	0.0074

5. Model Adequacy (Model 2)

For model adequacy checking on Model 2, we tested the following three assumptions:

1. Normality assumption of residuals ε_{ik}:
Residuals are assumed to be normally distributed at mean 0 and constant but unknown variance σ^2.

$$\varepsilon_{ik} \sim N(0, \sigma^2)$$

From output (Refer to Output -- page 14-16), we can see the residuals are distributed at mean 0 and variance 0.008. This distributed has a little skewness. Shipero-Wilk test gives a pretty large p-value: 0.9958. The distribution of residuals is not different from normal distribution. We can conclude residuals are normally distributed with mean 0 and variance 0.008. Stem leaf plot, box plot and normal probability plot support this result.

2. Independence assumption.
We used plot of residuals versus order: Res*order to test this assumption. We can see there's no clear pattern in this plot. The assumption of independence is not violated (Refer to Output – page 19).

3. Constant variances assumption:
We used Barlett's test and plot of residuals versus predicted response value to test this assumption.
Statistical tests for equality of variance: Bartlett's test:

$$X_0^2 = 2.3026 \frac{q}{c}$$

$$q = (N-a)\log_{10} S_p^2 - \sum_{i-1}^{a} (n_i - 1) \log_{10} S_i^2$$

$$c = 1 + \frac{1}{3(a-1)} \left(\sum_{i-1}^{a} (n_i-1)^{-1} - (N-a)^{-1} \right)$$

$$S_p^2 = \frac{\sum_{i=1}^{a} (n_i - 1) S_i^2}{N-a}$$

X_0^2=3.0357 with 2 df. P-value=0.22

We cannot reject the null hypothesis and conclude that all three variances are the same. The plot of residuals versus expected y is structureless (Refer to Output – page 17). This is the same conclusion reached by statistic analysis. The assumption of homogeneity of variance is not violated.

Based on these test results, we can conclude the model 2 is correct and assumptions of the model are satisfied.

6. Conclusion

Model 1:

1. The hypotheses of interest:

(1) *H_0 : No difference in response due to factor A (Brand).*
$H_0 : t_1 = t_2 = t_3$

(2) *H_0 : No difference in response due to factor B (Printing).*
$H_0 : b_1 = b_2$

(3) *H_0 : No effect of interaction of A and B.*
$H_0 : (tb)_{ij} = 0, \; i = 1, 2, 3; \; j = 1, 2$

(4) *H_0 : No covariate effect.*
$H_0 : g = 0$

2. Conclusion:

By the Output of SAS, we get that

- Covariate effect is significant (P-value=0.0431) .
- A*B is significant (P-value=0.0255).
- B is significant only within A_3 (P-value=0.0115).
- A is significant within B_1 and B_2.

We are interested in factor A and give the multiple comparisons of A. By the output of the multiple comparisons, we get that the effect of A_2 is different from those of A_1 and A_3. After estimating A_1 -A_2, A_1-A_3 and A_2-A_3, we also get that the effect of A_1 is different from those of A_1 and A_3.

Model 2:

The hypotheses of interest:

(1) H_0 : *No difference in response due to factor A (Brand).*
$H_0 : t_1 = t_2 = t_3$

(2) H_0 : *Equal slope.*
$H_0 : g_1 = g_2 = g_3$

3. **Conclusion:**

By the Output of SAS, we get that

- The property of equal slope is not significant (P-value of X*A is 0.3648) .
- Covariate effect is not significant (P-value=0.1917) .
- A is not significant (P-value of Type III = 0.0641), but it is significant by P-value of Type I which is 0.0222..

So, we can conclude that factor A is not significant for the response—Absorbency, but it is significant if we don't consider the effect of covariate.

7. Alternative Design

- Different Slopes

If we allow different slopes in Model 1, we will get extremely small p-values, which indicate that every effect is significant (Refer to the table below and Output – page 21-22). The reason is that we only have 12 data, while we have 11 parameters when we allow different slopes in Model 1.

	SS	d.f.	MS	F	p
A	0.06204697	2	0.03102349	979.61	0.0010
B	0.00184310	1	0.00184310	58.20	0.0168
A*B	0.07232635	2	0.03616317	1141.90	0.0009
X	0.00421449	1	0.00421449	133.08	0.0074
A*X	0.02371772	2	0.00421449	374.46	0.0027
A*X	0.00198326	1	0.00198326	62.62	0.0156
Error	0.00006334	2	0.00003167		

- Blocking

It would not be an alternative because we are using rate as a covariate here. The reason is that we cannot measure RATE before experiment. So we cannot block on Rate first.