

A Funnel Swirling Experiment Regression Application

Names

--	--

This experimental investigation will apply least squares regression line to summarize the nature of the dependence of a response variable (Y) on an explanatory variable (x).

Description of the Experiment

Consider the following experiment: A ball bearing will be rolled down a channel and into a funnel; it will swirl around in the funnel for a while and strike a bolt at the bottom of the funnel. The goal is to describe how, if at all, the swirling time depends on the distance up the channel at which the ball bearing is released.

- 1) The experiment consists of releasing a ball bearing from a pre-set location and recording the time it takes to swirl in the funnel and finally strike the bolt at the bottom of the funnel. The elapsed time from entering the funnel to striking the bolt should be recorded as accurately as possible using the stopwatch provided.
- 2) Your experiment will investigate the effect of changing the height (x) up the channel at which the ball is released (the "explanatory" variable) on (Y) the total swirling time (the "response" variable).
- 3) Each pair will complete a total of 18 repetitions of this experiment: 2 repetitions at each of 9 height settings.

Set-up

Each station should have the following equipment:

- Stopwatch
- Funnel and channel apparatus
- Ball bearing
- One-meter stick

The heights up the channel you will use for the experiment will vary from 12 inches to 36 inches. Before carrying out any replications of your experiment, you should ensure that angle of the channel is not so steep that the ball bearing will swirl out of the funnel due to excessive speed. Check this by carrying out a couple of test runs releasing the ball bearing at the maximum height setting, 36 inches from the end of the channel. If the ball bearing swirls out, you will have to use the clamp to lower

the high end of the channel until the ball bearing swirls properly to the bottom of the funnel. (Note: Refer to the demonstration for further details.)

Design of the Experiment

As noted above, for each of the height settings your group will record the swirling time two times. Use R to randomly determine the order of the trials for your experiment and record the order below (click [here](#) to see how to do this in R).

Height (x) in inches	Trial 1 Order	Trial 2 Order
12		
15		
18		
21		
24		
27		
30		
33		
36		

Data Collection

Now that you have determined the order for the 18 trials, you should proceed to carry out the experiment. Carefully complete the data worksheet below, filling in the swirling times for each of your trials. Except for the height up the channel, all other conditions should be held as constant as possible. For example, you should keep the funnel base in the same position each time and be sure that the channel stays in the notched groove of the funnel. If you experience a trial that is obviously "defective" (e.g., ball flies out of the funnel, funnel falls over, etc.), you may elect to repeat that trial and omit the "defective" time.

Funnel Swirling Data Worksheet		
Height (x)	Trial	Time (y) (sec)
12	1	
12	2	
15	1	
15	2	
18	1	
18	2	
21	1	
21	2	
24	1	
24	2	
27	1	
27	2	
30	1	
30	2	
33	1	
33	2	
36	1	
36	2	

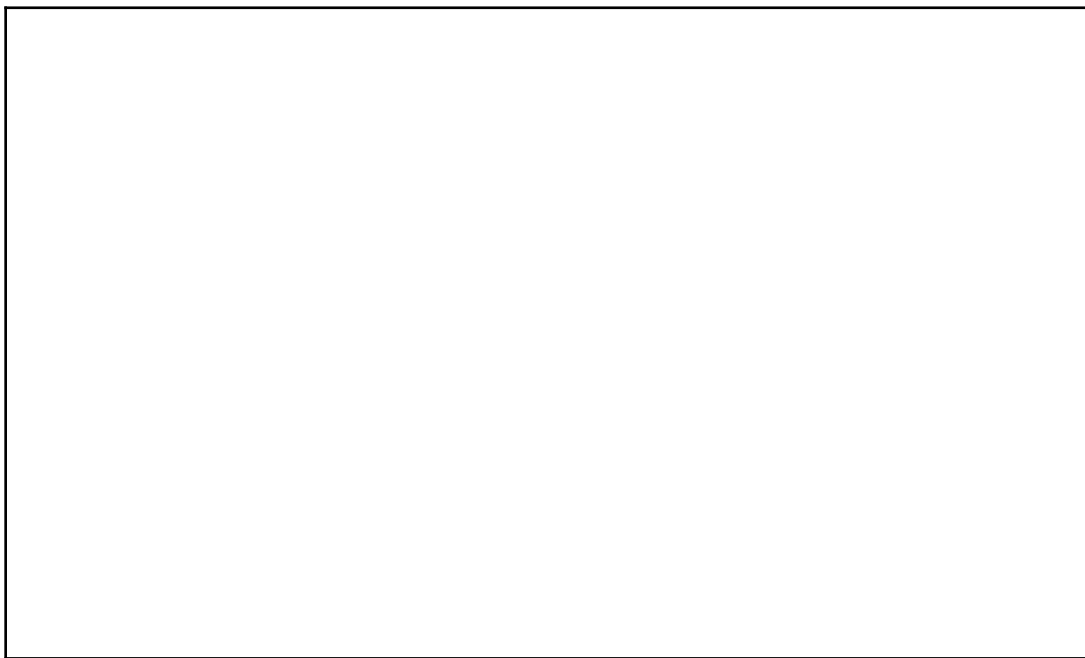
Data Entry

Enter your data into two vectors in R: `hght` (Channel height in inches) and `tswirl` (swirling time in seconds).

Data Display

- 1) On the axes below, sketch a scatterplot of `tswirl` (y) against `hght` (x). First do the plot in R using `plot(hght, tswirl)`.

Funnel Swirling Scatterplot



- 2) Describe the relationship between swirling time and height that you observe in your scatterplot. Comment as appropriate on the form, direction, and strength of the relationship and identify any outliers you see.
- 3) Using R^1 , compute the correlation coefficient, r , between `hght` and `tswirl`. Is the value of r consistent with the scatterplot?² Explain briefly.

¹ Compute the correlation coefficient r using the command `cor(hght, tswirl)`.

² A value of 1 or -1 for r means perfect linear association while a value of 0 means no linear association.

Least-Squares Regression Line

While the correlation coefficient indicates the **strength** of the linear association between two variables, the regression line specifies the **nature of the linear association**. The **regression line** is a straight line that describes how a response variable y changes as an explanatory variable x changes".

Computation of the estimated slope ($\hat{\beta}_1$) and intercept ($\hat{\beta}_0$) of the estimated regression line ($\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$) identifies the equation of the "best" line (the least squares line) passing through the cloud of points in the scatterplot. Use R to compute the equation of the least squares regression line that describes how swirling time (y) depends on height (x) using the following procedure. Click [here](#) to see how to do this in R.

- 1) Record the estimated regression equation in the space below (using the variable names rather than x and y).
- 2) Interpret the slope $\hat{\beta}_1$ of your regression equation by filling in the blank in the following sentence:

For every 1 inch increase in channel height, the funnel swirling time tends to increase by _____ seconds on average.
--

- 3) Now use R to add the regression line to your scatterplot. After completing this procedure for the lists `hght` and `tswirl`, add the least squares regression line to your scatterplot on page 4.

- 4) What swirling time does the least squares regression line predict for a ball released 34 inches up the channel? Record the prediction in the table below, and repeat for heights of 35 and 36 inches.³

Height (hght)	34 in	35 in	36in
Predicted Time (\hat{y})			

- 5) Do you notice a pattern in the predictions? By how many seconds is each prediction higher than the preceding one? Does this number look familiar (from your earlier calculations)? Explain.
- 6) Now carry out a t-test of significance to determine whether these data provide evidence at the $\alpha=0.05$ level that the true regression slope β_1 is not 0. Also use t-procedures to obtain a confidence interval for β_1 . The information needed for this is provided as part of the `summary()` output for a linear model.

³ You can get predictions using R by using the command `predict(fit, newdata=data.frame(x=c(34, 35, 36)))`. You should get the same answers as you get by plugging the values into your regression equation. To get more information on this command, lookup `predict.lm` in R. The command `predict` is a general "method" in R that behaves differently for different objects such as a linear model or `lm` object..

R Notes:

1. Random permutations

To obtain a random permutation, i.e., random ordering of an object x (such as a vector) use the command `sample(x)`. For example, to get a random permutation of the three numbers 2,4,10, use:

```
> x<-c(2,4,10)
> sample(x)
[1] 10  4  2
> sample(x)
[1]  2 10  4
>
```

Note that a new random order is obtained each time you submit the command.

Click [here](#) to return to the document.

2. Fitting a Least Squares Regression

To fit the least squares regression line for predicting y from x , use the R command `lm()` as follows:

```
> lm(y ~ x) # Fit the line and display the coefficients
```

Typically in R, we prefer to store the fitted line as an object so that other R commands can be used with it. For example, the following would store the fitted line under the name `funnelreg`. Subsequent commands `plot` and `summary` summarize the fit as noted in the comments. They may also be used to diagnose the quality of the fit of the regression.

```
> funnelreg<-lm(tswirl ~ hght) # Fit the line and save it
> plot(tswirl,hght) #Plot the data again
> abline(fit) #Add the regression line to the plot
> summary(fit) #Display summary statistics for the line
> coef(fit) #Display just the coefficients
> plot(fit) # Diagnostic plots
```

Click [here](#) to return.