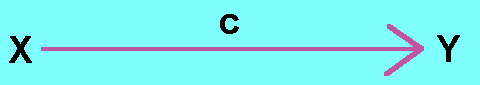
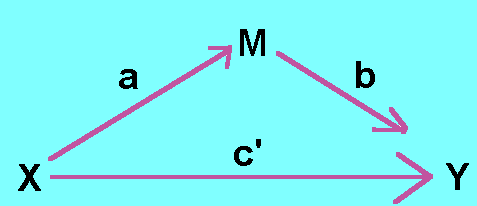
# COMPLETE EXAMPLE

# MEDIATION

# Please note that these steps are the traditional Baron and Kenny (1986) steps designed to instruct the basic idea of mediation. You can read more about mediation on Kenny’s website: <http://davidakenny.net/cm/mediate.htm>.





The steps:

1. Show that the X variable predicts Y (the *c* path).
2. Show that the X variable predicts the mediator M (the *a* path).
3. Show that M predicts Y controlling for X (the *b* path).
4. Show that X is lessened by including M in predicting Y (the *c’* path).
   1. Use a Sobel test for significance.

**Data set:** data 6.csv

**IV (X):**

* Baseball group – AL or NL

**Mediator (M):**

* Budget money for extra pitchers

**DV (Y):**

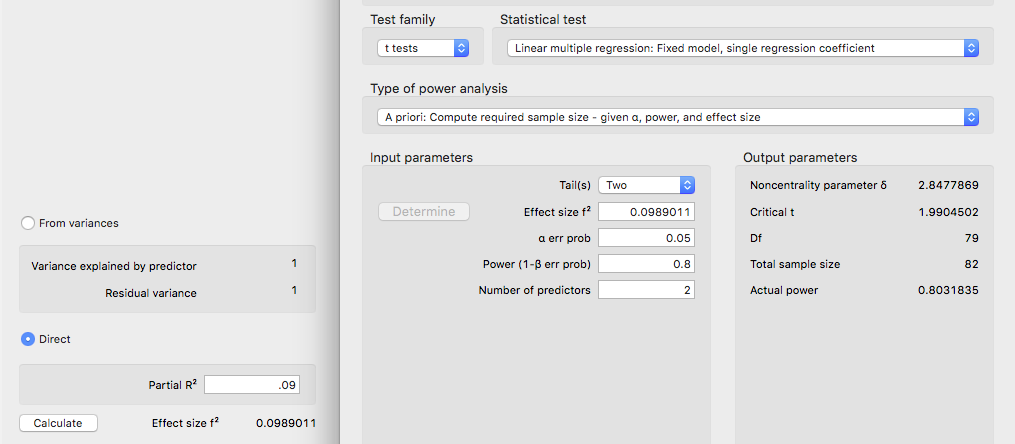
* Extra innings of pitching relief

**Research question:** Does the budget for pitchers mediate the relationship between baseball league and the innings of relief?

**Power:**

1. A tricky subject for mediation – you want X to originally predict Y, X to predict M, and X to stop predicting Y when M is included…
   1. The easiest guess for this type of test, is to estimate the effect size you’d expect for X to Y (originally), X to M, and M to Y.
   2. Others have written on this subject, but if you are truly using it for sample size estimation, then a best guess on those values will be adequate.
2. Open Gpower!
   1. Test family: t tests
   2. Statistical Test: Linear multiple regression: Fixed model, single regression coefficient
   3. Tails: two
   4. Effect size: click determine 🡪 direct 🡪 estimate partial R2 🡪 calculate and transfer to main window.
   5. Alpha = .05
   6. Power (1-beta of .20) = .80
   7. The number of predictors:
      1. For X 🡪 Y, you’d use 1
      2. For X 🡪 M, you’d use 1
      3. For M 🡪 Y, you’d use 2
      4. Pick the one you’d think has the lowest effect size since that will require the most people.
      5. For our examples, we will use M 🡪 Y.
3. Let’s estimate the following:
   1. Medium effect size (*R2* = .09)
   2. Number of predictors = 2 (X and M)

We would need 82 people to find a medium effect size of M to Y.



**Step 1: Does X predict Y?**

**C path**

## Linear Regression

| **Model Summary** | | | | | | | | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | | **R** | | **R²** | | **Adjusted R²** | | **RMSE** | | | **R² Change** | | **F Change** | | | | **df1** | | **df2** | | **p** | |
| 1 |  | 0.248 |  | 0.062 |  | 0.053 |  | 12.756 | |  | 0.062 |  | 7.028 | | |  | 1 |  | 107 |  | 0.009 |  |
|  | | | | | | | | | | | | | | | | | | | | | | |
| **ANOVA** | | | | | | | | | | | | | | | | | | | |
| **Model** | |  | | | **Sum of Squares** | | | | **df** | | **Mean Square** | | | **F** | | **p** | | | |
| 1 |  | Regression | |  | 1143 | | |  | 1 |  | 1143.5 | |  | 7.028 |  | 0.009 | | |  |
|  |  | Residual | |  | 17410 | | |  | 107 |  | 162.7 | |  |  |  |  | | |  |
|  |  | Total | |  | 18553 | | |  | 108 |  |  | |  |  |  |  | | |  |
|  | | | | | | | | | | | | | | | | | | | |

| **Coefficients** | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | |  | | **Unstandardized** | | | | **Standard Error** | | **Standardized** | | **t** | | **p** | |
| 1 |  | intercept |  | 19.342 | | |  | 1.881 |  |  |  | 10.284 |  | < .001 |  |
|  |  | league |  | -6.558 | | |  | 2.474 |  | -0.248 |  | -2.651 |  | 0.009 |  |
|  | | | | | | | | | | | | | | | |
| **Part And Partial Correlations** | | | | | | | | |
| **Model** | |  | | | **Partial** | | **Part** | |
| 1 |  | league | |  | -0.248 |  | -0.248 |  |
|  | | | | | | | | |

***F*(1, 107) = 7.03, *p* = .01, *R²* = .06**

***b* = -6.56, *t*(107) = -2.65, *p* = .01, *pr*2 = .06**

**AL has more innings pitching relief than NL.**

**Step 2: Does X predict M?**

**Linear Regression**

| **Model Summary** | | | | | | | | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | | **R** | | **R²** | | **Adjusted R²** | | **RMSE** | | | **R² Change** | | **F Change** | | | | **df1** | | **df2** | | **p** | |
| 1 |  | 0.236 |  | 0.056 |  | 0.047 |  | 3.752 | |  | 0.056 |  | 6.333 | | |  | 1 |  | 107 |  | 0.013 |  |
|  | | | | | | | | | | | | | | | | | | | | | | |
| **ANOVA** | | | | | | | | | | | | | | | | | | | |
| **Model** | |  | | | **Sum of Squares** | | | | **df** | | **Mean Square** | | | **F** | | **p** | | | |
| 1 |  | Regression | |  | 89.15 | | |  | 1 |  | 89.15 | |  | 6.333 |  | 0.013 | | |  |
|  |  | Residual | |  | 1506.19 | | |  | 107 |  | 14.08 | |  |  |  |  | | |  |
|  |  | Total | |  | 1595.34 | | |  | 108 |  |  | |  |  |  |  | | |  |
|  | | | | | | | | | | | | | | | | | | | |

| **Coefficients** | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | |  | | **Unstandardized** | | | | **Standard Error** | | **Standardized** | | **t** | | **p** | |
| 1 |  | intercept |  | 4.520 | | |  | 0.553 |  |  |  | 8.171 |  | < .001 |  |
|  |  | league |  | -1.831 | | |  | 0.728 |  | -0.236 |  | -2.517 |  | 0.013 |  |
|  | | | | | | | | | | | | | | | |
| **Part And Partial Correlations** | | | | | | | | |
| **Model** | |  | | | **Partial** | | **Part** | |
| 1 |  | league | |  | -0.236 |  | -0.236 |  |
|  | | | | | | | | |

***F*(1, 107) = 6.33, *p* = .01, *R²* = .06**

***b* = -1.83, *t*(107) = -2.52, *p* = .01, *pr*2 = .06**

**AL has a higher budget than the NL.**

**Step 3: One analysis, two questions**

**Does M predict Y with X?**

**Does X *lessen or no longer predict* Y with M?**

**Linear Regression**

| **Model Summary** | | | | | | | | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | | **R** | | **R²** | | **Adjusted R²** | | **RMSE** | | | **R² Change** | | **F Change** | | | | **df1** | | **df2** | | **p** | |
| 1 |  | 0.469 |  | 0.220 |  | 0.206 |  | 11.682 | |  | 0.220 |  | 14.98 | | |  | 2 |  | 106 |  | < .001 |  |
|  | | | | | | | | | | | | | | | | | | | | | | |
| **ANOVA** | | | | | | | | | | | | | | | | | | | |
| **Model** | |  | | | **Sum of Squares** | | | | **df** | | **Mean Square** | | | **F** | | **p** | | | |
| 1 |  | Regression | |  | 4088 | | |  | 2 |  | 2043.9 | |  | 14.98 |  | < .001 | | |  |
|  |  | Residual | |  | 14465 | | |  | 106 |  | 136.5 | |  |  |  |  | | |  |
|  |  | Total | |  | 18553 | | |  | 108 |  |  | |  |  |  |  | | |  |
|  | | | | | | | | | | | | | | | | | | | |

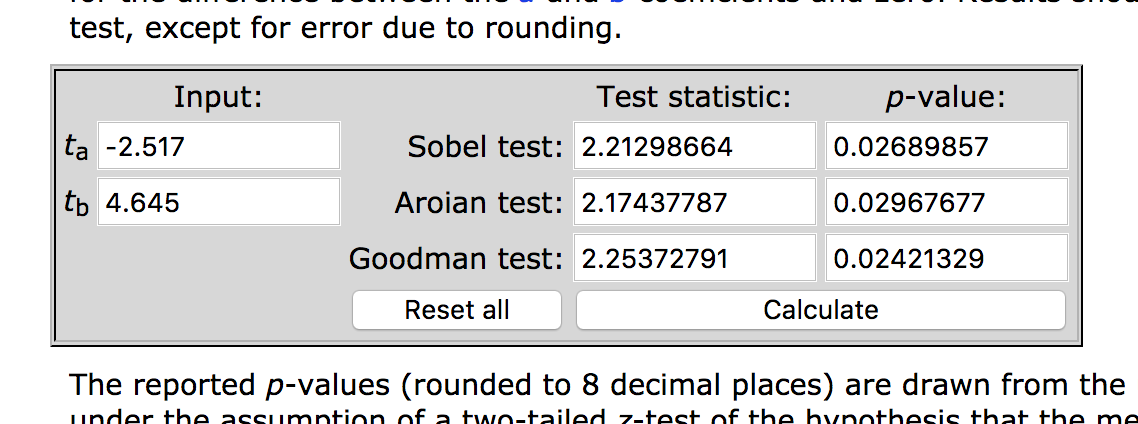
| **Coefficients** | | | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | |  | | **Unstandardized** | | | | **Standard Error** | | **Standardized** | | **t** | | **p** | |
| 1 |  | intercept |  | 13.023 | | |  | 2.195 |  |  |  | 5.933 |  | < .001 |  |
|  |  | league |  | -3.998 | | |  | 2.332 |  | -0.151 |  | -1.715 |  | 0.089 |  |
|  |  | budget |  | 1.398 | | |  | 0.301 |  | 0.410 |  | 4.645 |  | < .001 |  |
|  | | | | | | | | | | | | | | | |
| **Part And Partial Correlations** | | | | | | | | |
| **Model** | |  | | | **Partial** | | **Part** | |
| 1 |  | league | |  | -0.164 |  | -0.147 |  |
|  |  | budget | |  | 0.411 |  | 0.398 |  |
|  | | | | | | | | |

***F*(2, 106) = 14.98, *p* < .001, *R²* = .22**

**Budget (m predicting y, b path) *b* = 1.40, *t*(106) = 4.65, *p* < .001, *pr*2 = .17**

**League (x predicting y with m, c’ path) *b* = -4.00, *t*(106) = -1.72, *p* = .09, *pr*2 = .03**

1. The Sobel test:
   1. This test tells you if c and c’ are significant different. You could find that X predicts M which predicts Y but that c and c’ aren’t really changed by including M in the equation (which implies a lot of correlated variables, not mediation).
   2. Additionally, you want c’ to be *closer to zero* than c, indicating a decrease in the relationship. If c’ is greater or stronger than c, that indicates you might actually have an interactive effect, rather than a mediator (i.e. moderation).
   3. [**http://quantpsy.org/sobel/sobel.htm**](http://quantpsy.org/sobel/sobel.htm)



*Z* = 2.12, *p* = .03

**An example write up using the original dataset example before it became an awesome baseball example:**

**Results**

Treatment condition for housing (either treated or control group) was used to predict days in housing, with housing contacts expected to mediate the relationship between treatment condition and days in housing. Data were screened for outliers, and none were found. All other assumptions of regression were checked and appeared satisfactory.

See Figure 1 for visual diagram of the mediated relationship. First, using steps described by Baron and Kenny (1986), treatment was a significant predictor of days in housing (the *c* pathway), as shown in Table 1. The treatment condition showed a higher number of days in housing than the control condition, *t*(105) = 2.72, *p* = .01, *pr*2 = .06. Second, treatment condition was used to predict the mediator variable of housing contacts (the *a* pathway), which showed that treatment condition was positively related to housing contacts, *t*(105) = 2.98, *p* = .01, *pr*2 = .06. Third, the relationship between the mediator housing contacts and days in housing was examined controlling for the treatment condition (the *b* pathway). Number of housing contacts was positively related to the number of days in housing, *t*(104) = 4.96, *p* <.001, *pr*2 = .22. Lastly, the mediated relationship between treatment condition and days in housing was examined for a drop in prediction when the mediator was added to the model (the *c’* pathway). Mediation was found, showing that the relationship between treatment condition and days in housing was no longer significant after controlling for housing contacts, *t*(104) = 1.50, *p* = .14, *pr*2 = .02. The Sobel test was used to determine that the *ab* effect was significantly greater than zero, *Z* = 2.17, *p* = .02.

|  |
| --- |
| *b* 1.74  *a* 1.89  Housing Contacts  *c'* 3.55  *c* 6.82  Days in Housing  Treatment Condition |

*Figure 1.* Mediated relationship between treatment condition and days in housing with housing contacts as the mediator.

Table 1

*Model Summaries for Mediation Analysis*

|  |  |  |  |
| --- | --- | --- | --- |
| Model | *F* | *p* | *R2* |
| Treatment Condition predicting Days in Housing | (1, 105) = 7.38 | <.01 | .07 |
| Treatment Condition predicting Housing Contacts | (1, 105) = 8.87 | <.01 | .08 |
| Treatment Condition and Housing Contacts predicting Days in Housing | (1, 104) = 16.82 | <.001 | .24 |