4.04 Multiple regression: Individual tests

In this video you'll learn how to perform **individual t-tests** to see - for each predictor separately - whether it's significantly related to the response variable, while controlling for the other predictors.
Consider our example where we predicted popularity of cat videos - measured as number of page views–with the predictors cat age and hairiness - rated on a scale from zero to ten.
Suppose we performed an F test and found that - combined - the predictors are significantly related to video popularity. The obvious follow-up question is: Which of the predictors are responsible for this significant overall effect?

To answer this question we perform individual follow-up t-tests for each predictor to assess its relationship with the response variable, while controlling for the other predictors.

The assumptions that need to be met, in order for the t-tests to give valid results, are the same as for the overall test. These assumptions are: **linearity** of each predictor and the response variable, for each value of the other predictors, and **normality**, **homoscedasticity** and **independence** of the residuals. Also the number of observations needs to be high enough relative to the number of predictors. I'll discuss these assumptions and how to check them later on.

So have do we perform these individual tests? Well, the procedure is similar to simple linear regression. The null hypothesis states that for a particular predictor the regression coefficient equals zero, while controlling for the other predictors.
In regression we usually contrast the null hypothesis with a non-directional alternative hypothesis, stating that the regression coefficient does not equal zero. Directional alternative hypotheses are possible however.

The test statistic t equals the regression coefficient divided by its standard error: $t = \frac{b_j}{se_{b_j}}$. We'll leave the computation of the standard error to statistical software, since is difficult to compute when we control for the other predictors.
The degrees of freedom equal n, the number of observations, minus the number of parameters in our model, which equals the number of predictors plus one - for the intercept.

UNIVERSITY OF AMSTERDAM

Supposed that in our example the regression coefficient for cat age equals -1.775 and the standard error is 1.805. Then the t-value equals -0.983. With 5 - 3 = 2 degrees of freedom, we can now calculate or look up the p-value. The calculated two-sided p-value equals 0.429. If we use a table, we find that the right-sided p-value is at least smaller than 0.25, but larger than 0.10. So the two-sided p-value lies between 0.50 and 0.20. We can now conclude that cat age is not related to video popularity, while controlling for hairiness.

We can perform the same procedure for hairiness. Supposed the regression coefficient equals 1.414 and the standard error is 1.045. Then the t-value equals 1.353, with again 5 - 3 = 2 degrees of freedom.
The calculated two-sided p-value equals 0.309. If we use a table, we find that the right-sided p-value is at least smaller than 0.25, but - again - larger than 0.10. So the two-sided p-value is - again - between 0.50 and 0.20. We can now conclude that hairiness is not related to video popularity, while we control for age.

Of course, besides t-tests we can also calculate confidence intervals for the regression coefficients. The formula is the same as in simple regression. We calculate the boundaries of the interval by taking the regression coefficient and subtracting and adding the margin of error. The margin of error for a 95 percent confidence interval is the t-value - associated with n minus k degrees of freedom and half the significance level of 0.050 - times the standard error.

To obtain the 95 percent confidence interval for the predictor cat age we take the regression coefficient -1.775 and subtract and add the margin of error: Which is t - which equals 4.3027 for this sample size and two degrees of freedom - and we multiply it with the standard error 1.805. So the margin of error equals 7.766. This results in an interval that ranges from -9.541 to 5.991. This is a very wide interval that contains zero. This is not strange since there are only five observations!

To obtain the 95 percent confidence interval for hairiness we take the regression coefficient 1.414 and subtract and add t - which equals 4.3027 - times the standard error 1.045. This results in a margin of error of 4.496 and an interval that ranges from -3.082 to 5.910. Again, a very wide interval containing zero, which is not surprising given the low number of observations.

You may have noticed that the results for cat age differ from the results we obtained when we used simple regression with cat age as our only predictor. When we control for the influence of other predictors the

relation between cat age and video popularity can become stronger, because the noise from other variables is cancelled out. But it can also become weaker, if part of the relation is also explained by other variables.

Consider this Venn diagram showing the total variation in the response variable. The predictors cat age and hairiness each capture or explain a unique part of the response variable.
But they also show some overlap with each other. In standard multiple regression the individual t-tests assess the unique contribution of the predictors. The overall test assesses the overall association of the predictors with the response variable, so the overlap between predictors and the response variable *is* taken into account.