**TU Dublin City Campus**

**Probability and Statistical Inference**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Week 9 Practical**

**Step 1.** Download the zip file from Brightspace that contains the datafiles we are using this week. Unzip to a location that makes sense for you and make sure you know how to access the data from within RStudio.

**Step 2.** Review the markdown (and HTML output generated from this).

**Step 3:** Using the Survey.data dataset (dataset on well-being) build a logistic regression model to predict the probability of being in the category of smoking or not smoking (variable smoke 1 = yes; 2 = no) with the predictors of gender (variable sex) and total perceived stress (tpstress)

To make life easy – you can start with a single predictor

* Check the model against the assumptions of logistic regression.

**Step 4:** Extend your model by including another predictor if you get time.