Bivariate distributions：Every statistic that we focused on to this point have been descriptive statistics which described univariate distributions. Notice the prefix “uni”, meaning one. A univariate distribution tells the story of a single variable, while a bivariate distribution describes how two variables interrelate

The population mean: μ = E(Y)   
• The population variance: σ2 = V(Y)

Let’s assume that in the population, the values of Y are normally distributed. That is, we

say Yi ~ N(μ, σ2), where μ is the average value of Y and σ2 is the variance of Y, “~” means “distributed as”, and “N(.,.)” means “normal distribution”.

H0: μ = μ0 = 0 H1: μ ≠ 0 Test: t = (𝑌̅−0)/(𝑠/√100), Significance level (α): α = 0.05 |tα/2|, df

p-value is the probability of rejecting the null hypothesis when the null hypothesis is in fact true.

Table

Description automatically generated Type I: You reject H0 when H0 is true. (You find a significant difference only really by chance).   
   
Type II: You fail to reject H0 when H1 is true. (You don’t find a significant difference when there really is one).

In simple regression, in the population we can relate a variable X and an outcome Y via the model, 𝑌𝑖 =𝛼+𝛽𝑋𝑖 +𝜀𝑖

α is the intercept. We can interpret this as the average value of Y when X = 0.

β is the slope, the effect of a 1-unit change in X on Y.

Ŷ =𝑎+𝑏𝑋𝑖 fixed part a = estimate of the average value of Yi when Xi = 0

b = estimate of the average change in Yi associated with a 1-unit increase in Xi.

R2 provides information on the proportion of total variation in Y that is explained by the model.

se = the amount of variation in Yi not explained for by the model/ on average, the model is incorrect by () point when predicting the sat verbal with the variables.

SE(b)The average distance that observed values deviate from the regression line

95%CI of β 95 out of 100 samples will have a confidence interval that contains the true slope of the regression line

P=0.24>0.05 there is not significant evidence to reject H0 or there is not evidence that there is a relationship between X1 and Y (once the other covariates have been taken into account.)

Multiple regression population models: Yi = α + β1X1i + β2X2i + ... + βpXpi + εi

In multiple regression the coefficient estimated is a partial coefficient. Holding constant, controlling for. In simple regression models they are not.

α is again the intercept. It is interpreted as the average value of Y when X1 = X2 = ...

= Xp = 0.

β1 is a slope. It is the effect of a 1-unit increase in X1 on Y, holding constant these other variables.

εi = Yi – E(Yi) is the residual for observation i.

X2 has the large effect on Y. X2 has the largest absolute value. **standardized coefficient**

For 1 sd increase in teacher pay, Sat verbal increase 0.1 sd units, controlling for PT.

Omnibus tests: H0: β1 = β2 = ... = βk = 0 vs Ha: at least one βj ≠ 0. Incremental tests (A subset of slopes) H0: β1 = β2 = ... = βq = 0 (for q < p) vs Ha: at least one of βj ≠ 0 (j=1...q)

H0: β1= β2 = 0 Bj at least one not equal to 0.. Interpret the results. P=0.01<0.05. reject H0 and conclude that at least one of the covariates in the model explains variation in Y. F(2,48)=89.288

It is important to remember that the adjusted-R2 is useful when making predictions and when comparing models but is not useful as a descriptive of the model. For that, the R2 is better.

t2=F P=0.01<0.05, reject H0, the full model is Significantly better F(1,48)=127.9

ANOVA is applied to just two groups, and when therefore one can calculate both a t-statistic and an F-statistic from the same data. F=t^2.

Categorical variables: Yi = α + βXi + γ1Lefti + γ2Ambii +γ3Rurali + εi It’s additive because the it add effect of both the categorical and continuous variables

The regression lines for the two groups are always parallel. as the sample change the slope, intercept and residuals will change.

Simple regression only has one line. The slope, intercept and residuals will change with new samples in the simple regression. The adjusted R squared values are higher in the additive model so additive model fits the data better. If one variable is omitted, the simple regression line may not accurately reflect the relationship between the dependent variable and the independent variable.

α = he average value of Y when X = 0, for right-handers in urban locations

β = the effect of a 1-unit increase in X on Y, holding constant handedness and urbanicity

γ1 = the average difference in Y between left-handers and right-handers, holding constant X and urbanicity,

γ3 = the average difference in Y between rural and urban dwellers, holding constant X and handedness

This is called an additive model because the relationship between X and Y is the same for left and right-handers (i.e., parallel lines).

an incremental F-test Ho: y MW=y NE=y WE=0, Ha: at least one y≠0 Sig.F change <0.01, reject the null hypothesis. Regions are significant factor in accounting for the variation in Sat verbal

Hypothesis in (C) : Ho: y MW=0, Ha: y MW≠0Ho: y NE=0, Ha: y NE≠0Ho: y WE=0, Ha: y WE≠0

The test in c were test the hypothesis of each y individually. Incremental f- test test hypothesis with all y to see the effect of the factor, after accounting for X. so in incremental f- test regions explain in sat verbal after accounting for teacher pay. The test in c determine the differences in sat verbal those that are MW and SE holding constant teacher pay.

Interactions: Yi = α + βXi + γLefthandedi +δXi\*Lefthandedi + εi

Right: Yi = α + βXi + εi

Left: Yi = α + βXi + γ +δXi + εi = (α + γ) + (β + δ)Xi + εi

α = the average value of Y when X = 0 for right-handers

β = the effect of a 1-unit increase in X on Y for right-handers, main effect

γ = the average difference between Y for left- versus right-handers when X = 0, the average difference in sat math between MW and NE states is 110.9 when TP=0

δ = the effect of a 1-unit increase in X on the relationship between X and Y for left- versus right-handers = the difference in slopes between the relationship between X and Y for left- and right-handers, interaction effect

the average change in the relationship between TP and Sat math for MW as compared to NE states is

NE: Sat math = α + β TP + εi

MW: Sat math= α + β TP +y 1 + δ teacher = (α+y1) + (β+δ1) TP + εi

SE: Sat math= α + β TP +y SE + δ teacher= (α+y2) + (β+δ2) TP + εi

WE: Sat math= α + β TP +y WE + δ teacher = (α+y3) + (β+δ3) TP + εi

In the quadratic model, Yi = α + β1Xi + β2Xi2 + εi

α = the average value of Y when X = 0

β1 = the effect of a 1-unit increase in X on Y in the neighborhood of X = 0

β2 = the change in the effect of a 1-unit increase in X on Y as X increases

cubic model

β3 = the change in the change of the effect of a 1-unit increase of X on Y

Polynomial interaction model

Yi = α + β1Xc1i + β2Xc2i + β3Xc1i2 + β4Xc2i2 + β5Xc1i\*Xc2i + εi

As you might expect, interpreting parameters in this model requires careful thinking:

α = the average value of Y at the average value of X1 and X2

β1 = the effect of a 1-unit increase in X1 on Y in the neighborhood of 𝑋1̅̅̅ and 𝑋2̅̅̅

β2 = the effect of a 1-unit increase in X2 on Y in the neighborhood of 𝑋1̅̅̅ and 𝑋2̅̅̅

β3 = the effect of a 1-unit increase in X1 on the linear relationship between X1 & Y

β4 = the effect of a 1-unit increase in X2 on the linear relationship between X2 & Y

β5 = for a fixed value of X1, β5 is the effect of a 1-unit increase in X2 on slope of the

relationship between X1 & Y, OR

= for a fixed value of X2, β5 is the effect of a 1-unit increase in X1 on slope of the

relationship between X2 & Y.