201ab Quantitative methods
L.12 Linear model: Categorical predictors

GLM: Categorical predictors (factors)

- Why?
- How to use categorical predictors in R?
- Perspectives on categorical predictors.
- Coding categorical variables in regression.

- Variations that require extensions of LM
  - Unequal variance t-test or ANOVA
  - Repeated measures and other random effects / correlated error structures.

Why categorical predictors?

- Does mean y differ between…
  - Treatment and control?
  - Males and females?
  - Dogs and cats?

Predictor is treated as a dichotomous / binary categorical variable

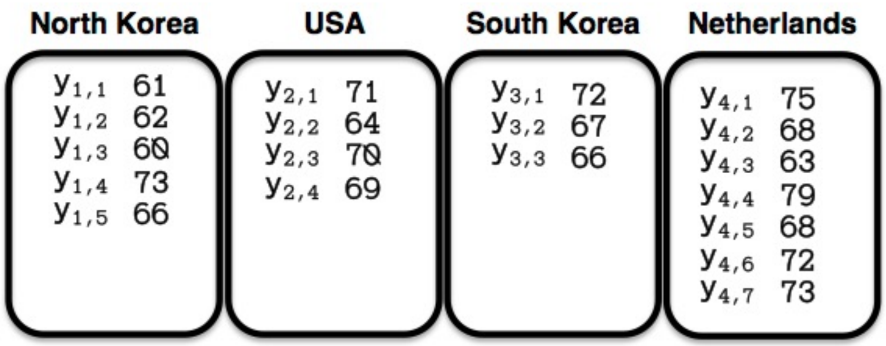
- Does mean y vary among…
  - Drug types?
  - Ethnicities? Religions? Etc.
  - Dog breeds?

Predictor is treated as a categorical variable

Do the groups have different means?

- If we have two groups, we can do a t-test.

- What if we have more than two groups?



- Lots of t-tests between pairs of groups are impractical, don't answer the right question.
- Instead we **test the variance of means across groups**: this is the "analysis of variance".

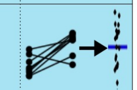
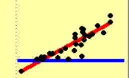
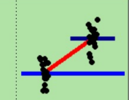
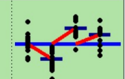
Overly specific named procedures

Response	~null	~binary	~category	~numerical	~numerical + category
Numerical	1-sample T-test	2-sample T-test	ANOVA	Regression, Pearson correlation	ANCOVA
Ranked-numerical		Mann-Whitney-U	Kruskall-Wallis	Spearman correlation	
2-category	Binomial test	Fisher's exact test	Chi-sq. indep.	Logistic regression	
k-category	Chi-sq. goodness of fit	Chi-squared independence			

Common statistical tests are linear models

Last updated: 28 June, 2019.Also check out the [Python version](#)!

See worked examples and more details at the accompanying notebook: https://lindeloev.github.io/tests-as-linear

	Common name	Built-in function in R	Equivalent linear model in R	Exact?	The linear model in words	Icon
Simple regression: lm(y ~ 1 + x)	**y is independent of x** P: One-sample t-test N: Wilcoxon signed-rank	t.test(y) wilcox.test(y)	lm(y ~ 1) lm(signed_rank(y) ~ 1)	✓ for N >14	One number (intercept, i.e., the mean) predicts **y**. - (Same, but it predicts the *signed rank* of **y**.)	
	P: Paired-sample t-test N: Wilcoxon matched pairs	t.test(y_1, y_2, paired=TRUE) wilcox.test(y_1, y_2, paired=TRUE)	lm(y_2 - y_1 ~ 1) lm(signed_rank(y_2 - y_1) ~ 1)	✓ for N >14	One intercept predicts the pairwise **y_2-y_1** differences. - (Same, but it predicts the *signed rank* of **y_2-y_1**.)	
	y ~ continuous x P: Pearson correlation N: Spearman correlation	cor.test(x, y, method='Pearson') cor.test(x, y, method='Spearman')	lm(y ~ 1 + x) lm(rank(y) ~ 1 + rank(x))	✓ for N >10	One intercept plus **x** multiplied by a number (slope) predicts **y**. - (Same, but with *ranked* **x** and **y**)	
	y ~ discrete x P: Two-sample t-test P: Welch's t-test N: Mann-Whitney U	t.test(y_1, y_2, var.equal=TRUE) t.test(y_1, y_2, var.equal=FALSE) wilcox.test(y_1, y_2)	lm(y ~ 1 + G_2)[A] gls(y ~ 1 + G_2, weights=...[B])[A] lm(signed_rank(y) ~ 1 + G_2)[A]	✓ ✓ for N >11	An intercept for **group 1** (plus a difference if **group 2**) predicts **y**. - (Same, but with one variance *per group* instead of one common.) - (Same, but it predicts the *signed rank* of **y**.)	
Multiple regression: lm(y ~ 1 + x_1 + x_2 +...)	P: One-way ANOVA N: Kruskal-Wallis	aov(y ~ group) kruskal.test(y ~ group)	lm(y ~ 1 + G_2 + G_3 +...+ G_N)[A] lm(rank(y) ~ 1 + G_2 + G_3 +...+ G_N)[A]	✓ for N >11	An intercept for **group 1** (plus a difference if group ≠ 1) predicts **y**. - (Same, but it predicts the *rank* of **y**.)	
	P: One-way ANCOVA	aov(y ~ group + x)	lm(y ~ 1 + G_2 + G_3 +...+ G_N + x)[A]	✓	- (Same, but plus a slope on **x**.) *Note: this is discrete AND continuous. ANCOVAs are ANOVAs with a continuous x.*	
	P: Two-way ANOVA	aov(y ~ group * sex)	lm(y ~ 1 + G_2 + G_3 +...+ G_N + S_2 + S_3 +...+ S_K + G_2*S_2 + G_3*S_3 + ... + G_N*S_K)	✓	Interaction term: changing **sex** changes the **y ~ group** parameters. *Note: **$G_{2 \text{ to } N}$** is an [indicator (0 or 1)](#) for each non-intercept levels of the **group** variable. Similarly for **$S_{2 \text{ to } K}$** for sex. The first line (with G_i) is main effect of group, the second (with S_i) for sex and the third is the **group × sex** interaction. For two levels (e.g. male/female), line 2 would just be "S_2" and line 3 would be S_2 multiplied with each G_i.*	*[Coming]*
	Counts ~ discrete x N: Chi-square test	chisq.test(groupXsex_table)	**Equivalent log-linear model** glm(y ~ 1 + G_2 + G_3 + ... + G_N + S_2 + S_3+ ... + S_K + G_2*S_2 + G_3*S_3 +...+ G_N*S_K, family=...)[A]	✓	Interaction term: (Same as Two-way ANOVA.) *Note: Run glm using the following arguments: `glm(model, family=poisson())` As linear-model, the Chi-square test is $\log(y_i) = \log(N) + \log(\alpha_i) + \log(\beta_j) + \log(\alpha_i\beta_j)$ where α_i and β_j are proportions. See more info in [the accompanying notebook](#).*	*Same as Two-way ANOVA*
	N: Goodness of fit	chisq.test(y)	glm(y ~ 1 + G_2 + G_3 +...+ G_N, family=...)[A]	✓	(Same as One-way ANOVA and see Chi-Square note.)	*1W-ANOVA*

List of common parametric (P) non-parametric (N) tests and equivalent linear models. The notation y ~ 1 + x is R shorthand for $y = 1 \cdot b + a \cdot x$ which most of us learned in school. Models in similar colors are highly similar, but really, notice how similar they *all* are across colors! For non-parametric models, the linear models are reasonable approximations for non-small sample sizes (see "Exact" column and click links to see simulations). Other less accurate approximations exist, e.g., Wilcoxon for the sign test and Goodness-of-fit for the binomial test. The signed rank function is `signed_rank = function(x) sign(x) * rank(abs(x))`. The variables G_i and S_i are ["dummy coded" indicator variables](#) (either 0 or 1) exploiting the fact that when Δx = 1 between categories the difference equals the slope. Subscripts (e.g., G_2 or y_1) indicate different columns in data. lm requires long-format data for all non-continuous models. All of this is exposed in greater detail and worked examples at https://lindeloev.github.io/tests-as-linear.

[A] See the note to the two-way ANOVA for explanation of the notation.
[B] Same model, but with one variance per group: `gls(value ~ 1 + G2, weights = varIdent(form = ~1|group), method="ML")`.

Jonas Kristoffer Lindeløv
https://lindeloev.net

Conceptually correct, but some restrictions apply.

Overly specific named procedures

Response	~null	~binary	~category	~numerical	~numerical + category
Numerical	1-sample T-test	2-sample T-test	ANOVA	Regression, Pearson correlation	ANCOVA
	`lm(y~1)`	`lm(y~f)`		`lm(y~x)`	`lm(y~x+f)`
Ranked-numerical		Mann-Whitney-U	Kruskall-Wallis	Spearman correlation	
		`~ lm(rank(y)~f)`		`~ lm(rank(y)~rank(x))`	
2-category	Binomial test	Fisher’s exact test	Chi-sq. indep.	Logistic regression	
	`glm(y~…, family=binomial())`				
k-category	Chi-sq. goodness of fit	Chi-squared independence			
	`~ glm(y~…, family=poisson())`				

Overly specific named procedures

Response	~null	~binary	~category	~numerical	~numerical + category
Numerical	1-sample T-test	2-sample T-test	ANOVA	Regression, Pearson correlation	ANCOVA
	`lm(y~1)`	`lm(y~f)`		`lm(y~x)`	`lm(y~x+f)`
Ranked-numerical		Mann-Whitney-U	Kruskall-Wallis	Spearman correlation	
		`~ lm(rank(y)~f)`		`~ lm(rank(y)~rank(x))`	
2-category	Binomial test	Fisher's exact test	Chi-sq. indep.	Logistic regression	
	`glm(y~…, family=binomial())`				
k-category	Chi-sq. goodness of fit	Chi-squared independence			
	`~ glm(y~…, family=poisson())`				

```
> grit = read_csv('http://vulstats.ucsd.edu/data/duckworth-grit-scale-data/data-coded.csv')

── Column specification ────────────────────────────────────────────────────────
cols(
  .default = col_double(),
  country = col_character(),
  gender = col_character(),
  hand = col_character(),
  race = col_character(),
  voted = col_character(),
  married = col_character(),
  operatingsystem = col_character(),
  browser = col_character()
)
ℹ Use `spec()` for the full column specifications.

> glimpse(grit)
Rows: 4,270
Columns: 27
$ country           <chr> "RO", "US", "US", "KE", "JP", "AU", "US", "RO", "EU", "NZ", "A…
$ surveyelapse      <dbl> 174, 120, 99, 5098, 340, 515, 126, 208, 130, 129, 592, 217, 26…
$ education         <dbl> 4, 2, 1, 3, 4, 3, 3, 2, 3, 1, 3, 2, 3, 2, 2, 1, 3, 3, 2, 4, 2,…
$ urban             <dbl> 3, 3, 2, 2, 2, 3, 2, 1, 3, 2, 1, 2, 3, 3, 3, 3, 3, 2, 3, 1, 3,…
$ gender            <chr> "female", "female", "female", "female", "male", "female", "mal…
$ engnat            <dbl> 2, 1, 2, 1, 2, 2, 1, 2, 2, 1, 2, 1, 1, 2, 2, 1, 2, 2, 2, 2, 1,…
$ age               <dbl> 28, 19, 16, 30, 38, 23, 35, 22, 50, 16, 52, 20, 23, 20, 23, 17…
$ hand              <chr> "right", "right", "right", "right", "right", "right", "right",…
$ religion          <dbl> 1, 6, 0, 6, 2, 12, 3, 7, 12, 1, 8, 12, 4, 10, 6, 12, 12, 2, 10…
$ orientation       <dbl> 1, 1, 1, 1, 1, 1, 1, 3, 1, 1, 1, 1, 1, 1, 1, 2, 1, 1, 4, 2, 1,…
$ race              <chr> "white or indigenous", "white or indigenous", "asian", "black"…
$ voted             <chr> "yes", "no", "no", "yes", "no", "no", "no", "no", "no", "no", …
$ married           <chr> "never", "never", "never", "never", "currently", NA, "previous…
$ familysize        <dbl> 2, 3, 3, 6, 3, 1, 1, 2, 3, 2, 3, 1, 3, 9, 3, 3, 1, 3, 2, 0, 1,…
$ operatingsystem   <chr> "Windows", "Macintosh", "Windows", "Windows", "Windows", "Wind…
$ browser           <chr> "Chrome", "Chrome", "Firefox", "Chrome", "Firefox", "Chrome", …
$ screenw           <dbl> 1366, 1280, 1920, 1600, 1920, 1920, 1366, 1366, 1600, 1440, 12…
$ screenh           <dbl> 768, 800, 1080, 900, 1080, 1080, 768, 768, 1000, 900, 1024, 76…
$ introelapse       <dbl> 69590, 33657, 95550, 4, 3, 2090, 36, 6, 14, 68, 726, 376, 3, 3…
$ testelapse        <dbl> 307, 134, 138, 4440, 337, 554, 212, 207, 183, 143, 311, 407, 8…
$ extroversion      <dbl> 1, 10, -12, -11, -18, 12, 10, 0, 14, 11, 0, -10, 0, -1, 4, -13…
$ neuroticism       <dbl> 18, 30, 23, 6, 23, 2, 28, 32, 3, 20, 2, 37, 13, 17, 27, 25, 2,…
$ agreeableness     <dbl> 19, 15, 9, 20, 9, 18, 12, 13, 23, 23, 12, 10, 20, 11, 11, 3, 1…
$ conscientiousness <dbl> 4, 11, 10, 20, 14, 18, 10, 18, 16, 10, 14, 15, 13, 7, -7, 6, 1…
$ openness          <dbl> 26, 24, 23, 22, 12, 28, 32, 17, 25, 22, 16, 26, 22, 25, 15, 10…
$ grit              <dbl> 0, -5, -3, -16, -1, -11, 5, 6, -15, -8, -2, 12, -15, 11, 11, 1…
$ vocabulary        <dbl> 10, 6, 11, 8, 4, 6, 13, 6, 12, 9, 6, 7, 9, 8, 6, 7, 5, 12, 2, …
```

GLM: 1-sample t-test

- Does the mean of a group differ from some null mean?
- E.g., does the mean level of *conscientiousness* deviate from random responses.
  - 10 (1-5 likert items), 6 positively coded, 4 negatively coded.
  - Mean expected from random responding: 6 (3*6 – 3*4)

GLM: 1-sample t-test

- Does the mean of a group differ from some null mean?
- E.g., does the mean level of *conscientiousness* deviate from random responses.
  - 10 (1-5 likert items), 6 positively coded, 4 negatively coded.
  - Mean expected from random responding: 6 = (3*6 – 3*4)

Via lm()

```
> lm(data = grit, (conscientiousness-6) ~ 1) %>% summary()

Call:
lm(formula = (conscientiousness - 6) ~ 1, data = grit)

Residuals:
    Min      1Q  Median      3Q     Max 
-23.713  -5.713   0.287   5.287  17.287 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   3.7130     0.1202   30.88   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.423 on 3811 degrees of freedom
```

Via t-test function

```
> t.test(x = grit$conscientiousness, mu = 6)

	One Sample t-test

data:  grit$conscientiousness
t = 30.883, df = 3811, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 6
95 percent confidence interval:
 9.477293 9.948730
sample estimates:
mean of x 
 9.713012 
```

GLM: 2-sample t-test

- Do the two groups have the same mean?
- E.g., does the mean level of *conscientiousness* differ between males and females?

GLM: 2-sample t-test

- Do the two groups have the same mean?
- E.g., does the mean level of *conscientiousness* differ between males and females?

Via lm()

```
> lm(data = grit, conscientiousness ~ gender) %>% summary()

Call:
lm(formula = conscientiousness ~ gender, data = grit)

Residuals:
     Min       1Q   Median       3Q      Max
-23.6800  -5.6800   0.2694   5.3200  17.2694

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   9.7306     0.1489   65.36   <2e-16 ***
gendermale   -0.0506     0.2525   -0.20    0.841
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.424 on 3810 degrees of freedom
Multiple R-squared:  1.054e-05, Adjusted R-squared:  -0.0002519
F-statistic: 0.04016 on 1 and 3810 DF,  p-value: 0.8412
```

Via t-test function

```
> t.test(grit$conscientiousness[grit$gender == 'male'],
+        grit$conscientiousness[grit$gender == 'female'],
+        var.equal=T)

        Two Sample t-test

data:  grit$conscientiousness[grit$gender == "male"] and g
ess[grit$gender == "female"]
t = -0.20039, df = 3810, p-value = 0.8412
alternative hypothesis: true difference in means is not ec
95 percent confidence interval:
 -0.5456563  0.4444581
sample estimates:
mean of x mean of y
 9.680000  9.730599
```

Do the groups have different means?

- If we have 1 group and a point null for mean, we test the intercept: lm(y~1) -- a “one-sample t-test”

- If we have 2 groups and a null of same means: we test the difference coef: lm(y~f) -- a “2-sample t-test”.

- If we have 3+ groups and a null of same means: we test the ANOVA: lm(y~f) – an “analysis of variance”
  - Lots of t-tests between pairs of groups are impractical, don’t answer the right question.
  - Instead we **test the variance of means across groups**: this is the “analysis of variance”.

GLM: one-way anova

- Do the groups have the same mean?
  i.e., is there non-zero variance across group means?
- E.g., does the mean level of *conscientiousness* differ among religions?

GLM: one-way anova

- Do groups have same mean? Variance across group means?
- does mean *conscientiousness* differ among religions?

```
> lm(data = grit, conscientiousness ~ religion) %>% summary()

Call:
lm(formula = conscientiousness ~ religion, data = grit)

Residuals:
    Min      1Q  Median      3Q     Max
-23.572  -5.057  -0.029   5.186  17.943

Coefficients:
                               Estimate Std. Error t value Pr(>|t
(Intercept)                      8.4517     0.2969  28.466  < 2e-1
religionAtheist                 -0.3950     0.4132  -0.956   0.3392
religionBuddhist                 0.3056     0.7817   0.391   0.6959
religionChristian (Catholic)     2.3623     0.3914   6.035 1.74e-09 ***
religionChristian (Mormon)       2.5727     1.1840   2.173   0.0298 *
religionChristian (Other)        2.5773     0.4128   6.244 4.74e-10 ***
religionChristian (Protestant)   1.8073     0.4577   3.949 7.99e-05 ***
religionHindu                    1.1205     0.6320   1.773   0.0763 .
religionJewish                   1.2379     1.0083   1.228   0.2196
religionMuslim                   0.7632     0.5695   1.340   0.1803
religionSikh                     2.2325     1.7096   1.306   0.1917
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.339 on 3801 degrees of freedom
Multiple R-squared:  0.02511,   Adjusted R-squared:  0.02255
F-statistic: 9.791 on 10 and 3801 DF,  p-value: 2.405e-16
```

```
> lm(data = grit, conscientiousness ~ religion) %>% anova()
Analysis of Variance Table

Response: conscientiousness
            Df Sum Sq Mean Sq F value    Pr(>F)
religion    10   5274  527.35  9.7913 2.405e-16 ***
Residuals 3801 204720   53.86
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

GLM: two-way anova

- Does mean vary across either/both factors? Consistently? does mean *conscientiousness* vary among religion, gender?

GLM: two-way anova

- Does mean vary across either/both factors? Consistently? does mean *conscientiousness* vary

```
> lm(data = grit, conscientiousness ~ gender*religion) %>% summary()

Call:
lm(formula = conscientiousness ~ gender * religion, data = grit)

Residuals:
     Min       1Q   Median       3Q      Max
-24.2125  -5.0684   0.0868   5.3670  18.1304

Coefficients:
                                          Estimate Std. Error t va
(Intercept)                                 8.9059     0.3695  24.
gendermale                                 -1.2728     0.6186  -2.
religionAtheist                            -0.8375     0.5379  -1.
religionBuddhist                           -0.2279     1.0227  -0.
religionChristian (Catholic)                1.7266     0.4777   3.
religionChristian (Mormon)                  3.4541     1.5109   2.
religionChristian (Other)                   1.7005     0.5007   3.
religionChristian (Protestant)              1.4872     0.5525   2.
religionHindu                               0.1157     0.8447   0.
religionJewish                              1.9799     1.2921   1.
religionMuslim                             -0.7670     0.7135  -1.075 0.282493
religionSikh                                1.4513     1.9922   0.728 0.466367
gendermale:religionAtheist                  1.2476     0.8448   1.477 0.139820
gendermale:religionBuddhist                 1.4585     1.5848   0.920 0.357460
gendermale:religionChristian (Catholic)     1.8914     0.8338   2.268 0.023356 *
gendermale:religionChristian (Mormon)      -2.1497     2.4253  -0.886 0.375488
gendermale:religionChristian (Other)        2.7691     0.8870   3.122 0.001811 **
gendermale:religionChristian (Protestant)   0.8004     0.9885   0.810 0.418185
gendermale:religionHindu                    2.4638     1.2768   1.930 0.053723 .
gendermale:religionJewish                  -1.7433     2.0612  -0.846 0.397721
gendermale:religionMuslim                   4.1935     1.1807   3.552 0.000387 ***
gendermale:religionSikh                     2.5157     3.8660   0.651 0.515267
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.325 on 3790 degrees of freedom
Multiple R-squared:  0.03164,	Adjusted R-squared:  0.02628
F-statistic: 5.897 on 21 and 3790 DF,  p-value: 2.859e-16
```

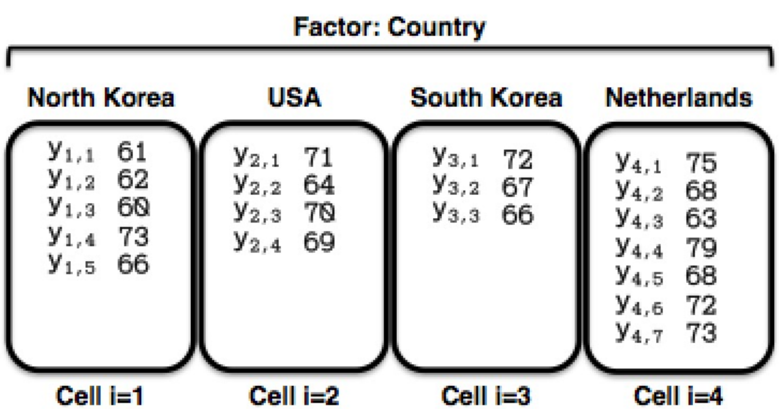
```
> lm(data = grit, conscientiousness ~ gender*religion) %>% anova()
Analysis of Variance Table

Response: conscientiousness
                  Df Sum Sq Mean Sq F value    Pr(>F)
gender             1      2    2.21  0.0412  0.839068
religion          10   5328  532.85  9.9311 < 2.2e-16 ***
gender:religion   10   1314  131.38  2.4487  0.006529 **
Residuals       3790 203350   53.65
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Three ways to think about factors

Cell organization:

Common formulation for doing ANOVA calculation by hand.

We avoid hand calculations, but this formulation helps understand what we are estimating.



Tidy data frame/table:

How we will see our data.

Data point	Country	Height
1	North Korea	61
2	North Korea	62
3	North Korea	60
4	North Korea	73
5	North Korea	66
6	USA	71
7	USA	64
8	USA	70
9	USA	69
10	South Korea	72
11	South Korea	67
12	South Korea	66
13	Netherlands	75
14	Netherlands	68
15	Netherlands	63
16	Netherlands	79
17	Netherlands	68
18	Netherlands	72
19	Netherlands	73

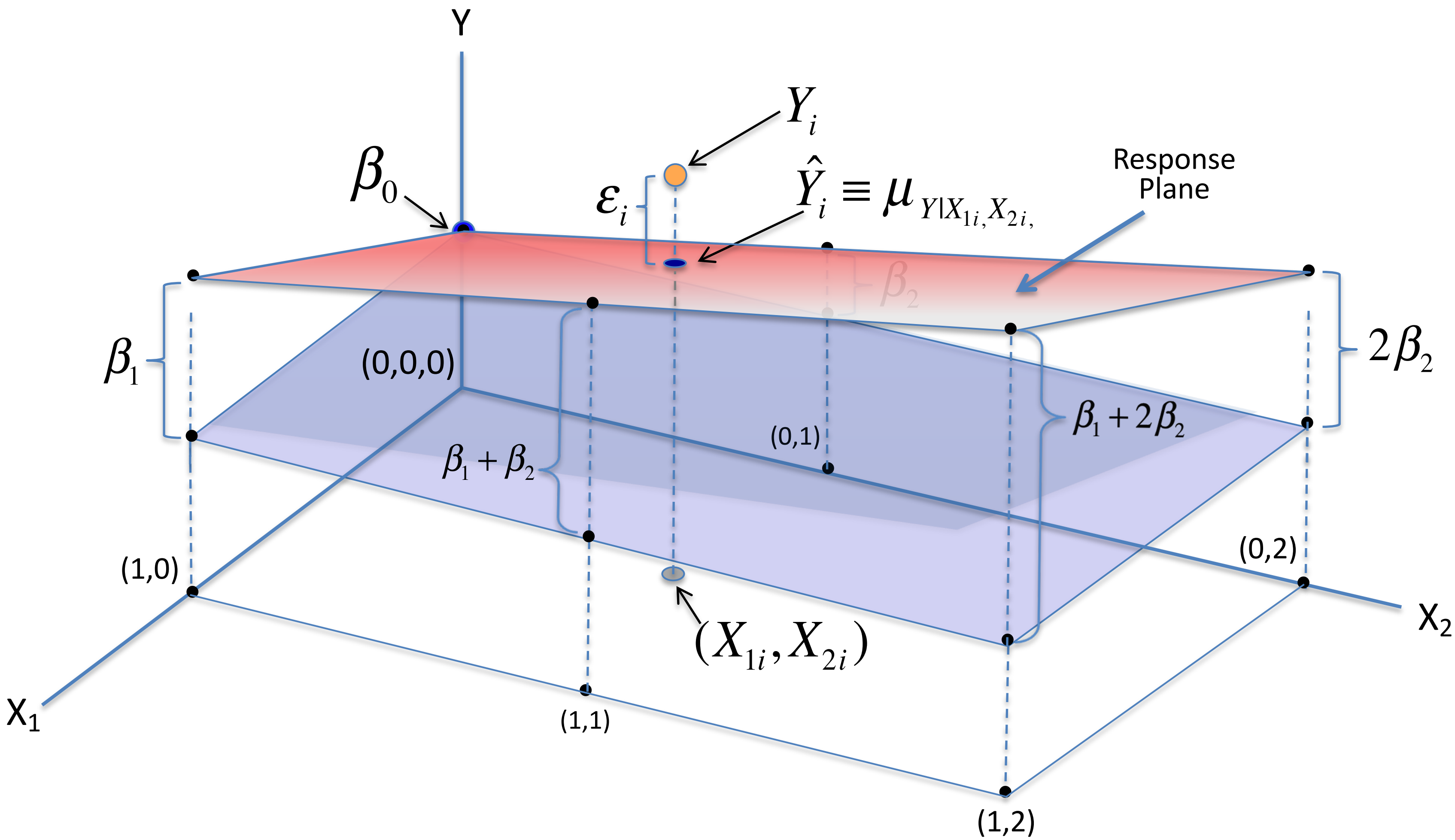
Matrix notation:

How statistical software represents our data to do the analysis.

Makes it easier to think about coding schemes.

$$
\begin{array}{c}
Y \\
\begin{bmatrix} 61 \\ 62 \\ 60 \\ 73 \\ 66 \\ 71 \\ 64 \\ 70 \\ 69 \\ 72 \\ 67 \\ 66 \\ 75 \\ 68 \\ 63 \\ 79 \\ 68 \\ 72 \\ 73 \end{bmatrix}
\end{array}
\qquad
\begin{array}{c}
\begin{array}{cccc} X1 & X2 & X3 & X4 \end{array} \\
\begin{pmatrix}
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1
\end{pmatrix}
\end{array}
$$

$$\boxed{Y_i} = \boxed{\beta_0} + \boxed{\beta_1}\boxed{X_{1i}} + \boxed{\beta_2}\boxed{X_{2i}} + \boxed{\varepsilon_i}$$



FROM JULIAN PARRIS

$$\boxed{Y_i} = \boxed{\beta_0} + \boxed{\beta_1}\boxed{X_{1i}} + \boxed{\beta_2}\boxed{X_{2i}} + \boxed{\varepsilon_i}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \ldots \\ y_i \\ \ldots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ 1 & x_{13} & x_{23} \\ \ldots & \ldots & \ldots \\ 1 & x_{1i} & x_{2i} \\ \ldots & \ldots & \ldots \\ 1 & x_{1n} & x_{2n} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \ldots \\ \varepsilon_i \\ \ldots \\ \varepsilon_n \end{bmatrix}$$

$$\boxed{Y_i} = \boxed{\beta_0} + \boxed{\beta_1}\boxed{X_{1i}} + \boxed{\beta_2}\boxed{X_{2i}} + \boxed{\varepsilon_i}$$

All the y data points in a single vector

$$\boxed{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \ldots \\ y_i \\ \ldots \\ y_n \end{bmatrix}} = \begin{bmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ 1 & x_{13} & x_{23} \\ \ldots & \ldots & \ldots \\ 1 & x_{1i} & x_{2i} \\ \ldots & \ldots & \ldots \\ 1 & x_{1n} & x_{2n} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \ldots \\ \varepsilon_i \\ \ldots \\ \varepsilon_n \end{bmatrix}$$

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$$

All of the x predictors in one matrix.
(constant 1 for the intercept: sometimes called X0)

All the y data points in a single vector

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \ldots \\ y_i \\ \ldots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ 1 & x_{13} & x_{23} \\ \ldots & \ldots & \ldots \\ 1 & x_{1i} & x_{2i} \\ \ldots & \ldots & \ldots \\ 1 & x_{1n} & x_{2n} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \ldots \\ \varepsilon_i \\ \ldots \\ \varepsilon_n \end{bmatrix}$$

$$\boxed{Y_i} = \boxed{\beta_0} \cdot \boxed{1} + \boxed{\beta_1}\boxed{X_{1i}} + \boxed{\beta_2}\boxed{X_{2i}} + \boxed{\varepsilon_i}$$

All of the x predictors in one matrix.
(constant 1 for the intercept: sometimes called X0)

All the y data points in a single vector

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \ldots \\ y_i \\ \ldots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ 1 & x_{13} & x_{23} \\ \ldots & \ldots & \ldots \\ 1 & x_{1i} & x_{2i} \\ \ldots & \ldots & \ldots \\ 1 & x_{1n} & x_{2n} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \ldots \\ \varepsilon_i \\ \ldots \\ \varepsilon_n \end{bmatrix}$$

$$\boxed{Y_i} = \boxed{\beta_0} + \boxed{\beta_1}\boxed{X_{1i}} + \boxed{\beta_2}\boxed{X_{2i}} + \boxed{\varepsilon_i}$$

All of the x predictors in one matrix.
(constant 1 for the intercept: sometimes called X0)

All the y data points in a single vector

All of the coefficients in a single vector

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_i \\ \dots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ 1 & x_{13} & x_{23} \\ \dots & \dots & \dots \\ 1 & x_{1i} & x_{2i} \\ \dots & \dots & \dots \\ 1 & x_{1n} & x_{2n} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \dots \\ \varepsilon_i \\ \dots \\ \varepsilon_n \end{bmatrix}$$

$$\boxed{Y_i} = \boxed{\beta_0} + \boxed{\beta_1}\boxed{X_{1i}} + \boxed{\beta_2}\boxed{X_{2i}} + \boxed{\varepsilon_i}$$

All of the x predictors in one matrix.
(constant 1 for the intercept: sometimes called X0)

All the y data points in a single vector

All of the coefficients in a single vector

All the errors (residuals) in a single vector

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_i \\ \dots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ 1 & x_{13} & x_{23} \\ \dots & \dots & \dots \\ 1 & x_{1i} & x_{2i} \\ \dots & \dots & \dots \\ 1 & x_{1n} & x_{2n} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \dots \\ \varepsilon_i \\ \dots \\ \varepsilon_n \end{bmatrix}$$

$$Y_i = \boxed{\beta_0} + \boxed{\beta_1} X_{1i} + \boxed{\beta_2} X_{2i} + \varepsilon_i$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_i \\ \dots \\ y_n \end{bmatrix} = \overbrace{\begin{bmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ 1 & x_{13} & x_{23} \\ \dots & \dots & \dots \\ 1 & x_{1i} & x_{2i} \\ \dots & \dots & \dots \\ 1 & x_{1n} & x_{2n} \end{bmatrix} \begin{bmatrix} \boxed{\beta_0} \\ \boxed{\beta_1} \\ \boxed{\beta_2} \end{bmatrix}}^{\text{This matrix multiplication yields an } n \text{ unit vector, each element of which is y.hat}_i\text{: B0*1 + B1*x}_{1i}\text{ + B2*x}_{2i}} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \dots \\ \varepsilon_i \\ \dots \\ \varepsilon_n \end{bmatrix}$$

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_i \\ \dots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ 1 & x_{13} & x_{23} \\ \dots & \dots & \dots \\ 1 & x_{1i} & x_{2i} \\ \dots & \dots & \dots \\ 1 & x_{1n} & x_{2n} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \dots \\ \varepsilon_i \\ \dots \\ \varepsilon_n \end{bmatrix}$$

- Matrix notation highlights…
  - …there is no qualitative difference between slopes and intercept.
  - …the design of various indicator variables.

The design matrix encodes variables for regression

Generally, this is something that R/SPSS/JMP does for us behind the scenes, and we don't need to worry about how the design matrix is set up. There are different acceptable/correct ways to do this coding, and a great many ways to do it very incorrectly.

Y	X1	X2	X3	X4
61	1	0	0	0
62	1	0	0	0
60	1	0	0	0
73	1	0	0	0
66	1	0	0	0
71	0	1	0	0
64	0	1	0	0
70	0	1	0	0
69	0	1	0	0
72	0	0	1	0
67	0	0	1	0
66	0	0	1	0
75	0	0	0	1
68	0	0	0	1
63	0	0	0	1
79	0	0	0	1
68	0	0	0	1
72	0	0	0	1
73	0	0	0	1

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \ldots \\ y_i \\ \ldots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 1 & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \ldots \\ \varepsilon_i \\ \ldots \\ \varepsilon_n \end{bmatrix}$$

Different coding schemes

$$
\begin{array}{r}
\\ \\ \\ \\ \text{Men} \\ \\ \\ \\ \\ \\ \\ \\ \\ \text{Women} \\ \\ \\ \\ \\
\end{array}
\begin{pmatrix}
1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\
1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1
\end{pmatrix}
\quad
\begin{pmatrix}
1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\
1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0
\end{pmatrix}
\quad
\begin{pmatrix}
0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\
1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0
\end{pmatrix}
\quad
\begin{pmatrix}
1 & -1 \\ 1 & -1 \\ 1 & -1 \\ 1 & -1 \\ 1 & -1 \\ 1 & -1 \\ 1 & -1 \\ 1 & -1 \\ 1 & -1 \\ 1 & -1 \\
1 & +1 \\ 1 & +1 \\ 1 & +1 \\ 1 & +1 \\ 1 & +1 \\ 1 & +1 \\ 1 & +1 \\ 1 & +1
\end{pmatrix}
$$

These (and other) categorical variable coding schemes can capture that men and women have different, non-zero means.

However, the interpretation of B_0 and B_1 is very different in these cases.

And the "significance" of the coefficients means different things.

Lots of different coding schemes…

Dummy: compare each level to reference level, intercept at first level (default in R).

Simple: compare each level to reference level, but intercept is at overall mean

Deviation: Contrast coding comparing each level (except last) to grand mean.

Orthogonal polynomial: breaks down effects of ordinal variables into linear, quadratic, etc. trends.

Helmert: compare each level to mean of subsequent levels. (or reverse Helmert: each to mean of previous levels)

Forward difference: compare each level to the next. (or Backward difference: each level to the previous)

- Default factor coding scheme varies with software
- They all capture the same sources of variation, but the coefficients mean different things.
  - We will consider these sorts of comparisons when we deal with contrasts, rather than altering R's default coding scheme.

Geometric thinking about coefficients

	height	weight	sex
1	70	121	m
2	78	256	m
3	69	153	m
4	68	168	m
5	70	147	m
6	68	213	m
7	65	91	m
8	72	212	m
9	66	135	m
10	73	191	m
11	60	101	f
12	62	131	f
13	69	152	f
14	66	184	f
15	63	88	f
16	65	147	f
17	63	122	f
18	63	97	f

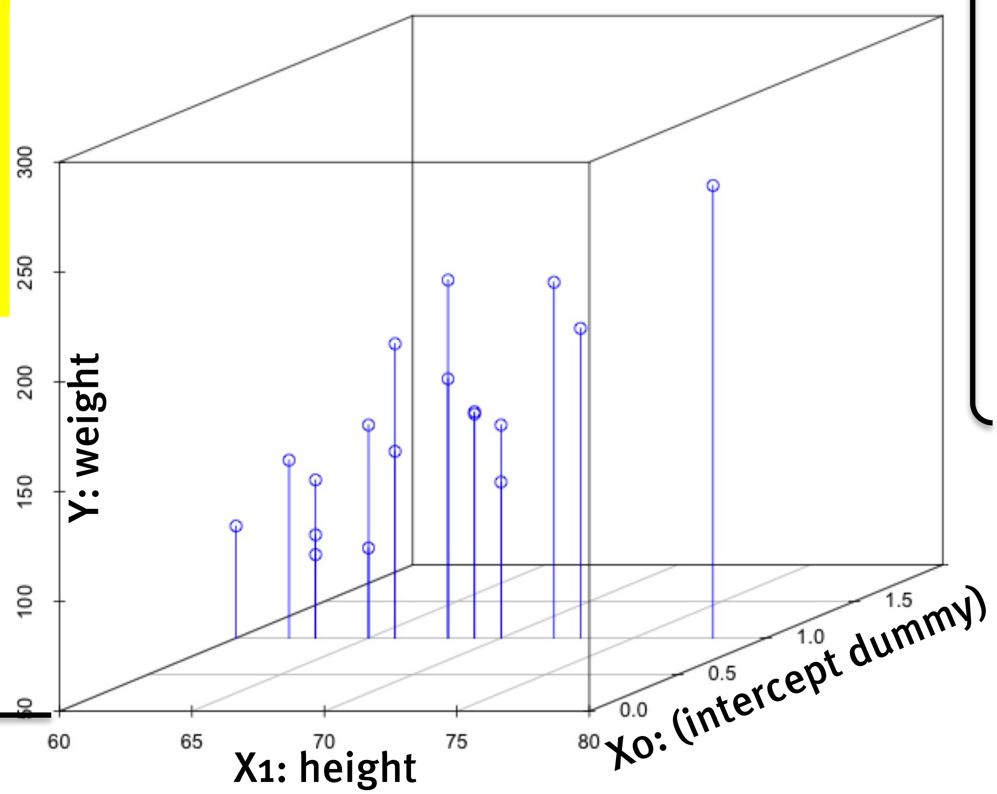
When we tell R to regress weight~height

Y: weight

$$\begin{pmatrix} 121 \\ 256 \\ 153 \\ 168 \\ 147 \\ 213 \\ 91 \\ 212 \\ 135 \\ 191 \\ 101 \\ 131 \\ 152 \\ 184 \\ 88 \\ 147 \\ 122 \\ 97 \end{pmatrix}$$

X: intercept + height

$$\begin{pmatrix} 1 & 70 \\ 1 & 78 \\ 1 & 69 \\ 1 & 68 \\ 1 & 70 \\ 1 & 68 \\ 1 & 65 \\ 1 & 72 \\ 1 & 66 \\ 1 & 73 \\ 1 & 60 \\ 1 & 62 \\ 1 & 69 \\ 1 & 66 \\ 1 & 63 \\ 1 & 65 \\ 1 & 63 \\ 1 & 63 \end{pmatrix}$$



Note: o has to be somehow represented. In this case, it is way over there.

Geometric thinking about coefficients

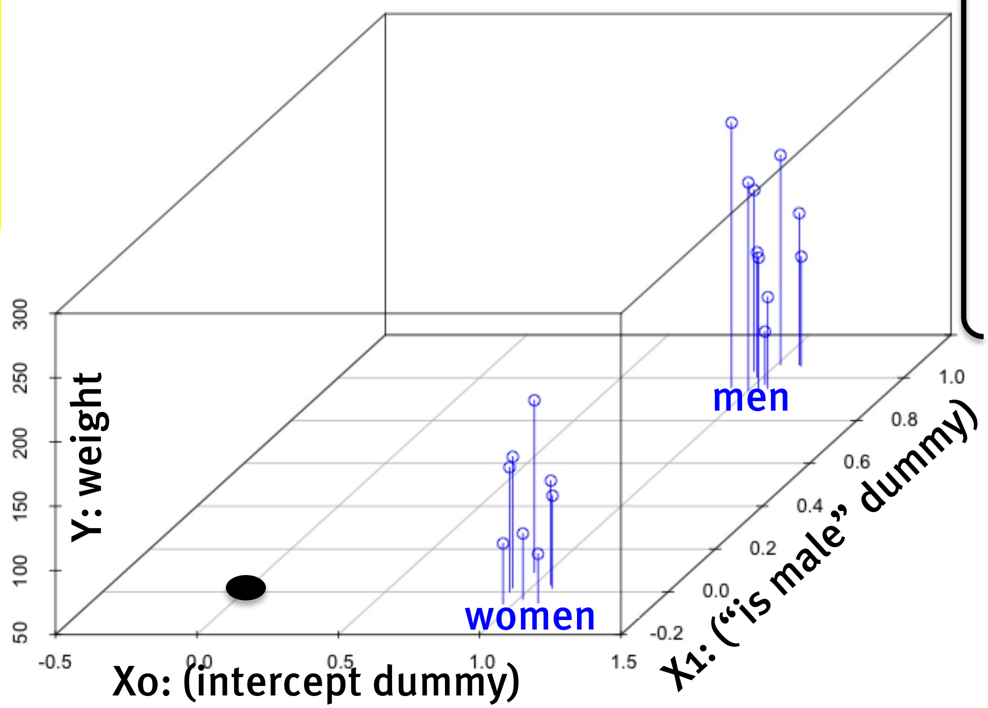
	height	weight	sex
1	70	121	m
2	78	256	m
3	69	153	m
4	68	168	m
5	70	147	m
6	68	213	m
7	65	91	m
8	72	212	m
9	66	135	m
10	73	191	m
11	60	101	f
12	62	131	f
13	69	152	f
14	66	184	f
15	63	88	f
16	65	147	f
17	63	122	f
18	63	97	f

When we tell R to regress weight~sex →

Y: weight **X: intercept + male?**

$$\begin{pmatrix} 121 \\ 256 \\ 153 \\ 168 \\ 147 \\ 213 \\ 91 \\ 212 \\ 135 \\ 191 \\ 101 \\ 131 \\ 152 \\ 184 \\ 88 \\ 147 \\ 122 \\ 97 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix}$$

So the average of women is captured by B_0.
The average of men is captured by B_0+B_1
B_1 = difference between avg men and women



Geometric thinking about coefficients

	height	weight	sex
1	70	121	m
2	78	256	m
3	69	153	m
4	68	168	m
5	70	147	m
6	68	213	m
7	65	91	m
8	72	212	m
9	66	135	m
10	73	191	m
11	60	101	f
12	62	131	f
13	69	152	f
14	66	184	f
15	63	88	f
16	65	147	f
17	63	122	f
18	63	97	f

An alternate way to code for gender.

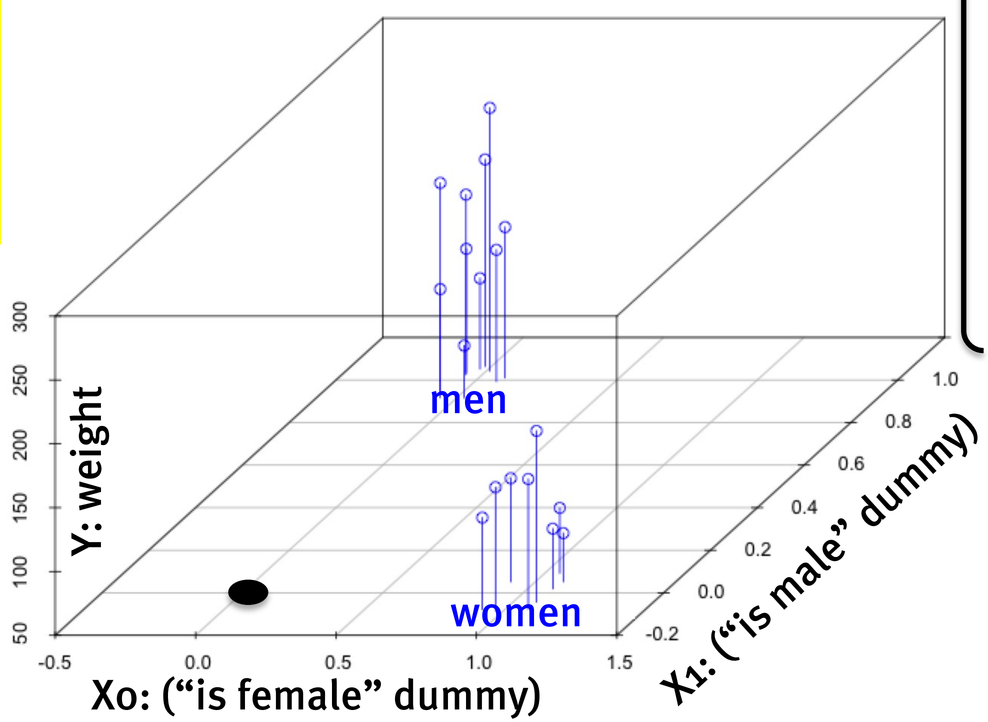
So the average of women is captured by B_0.
The average of men is captured by B_1
$B_0 - B_1$ = difference between avg men and women

Y: weight

$$\begin{pmatrix} 121 \\ 256 \\ 153 \\ 168 \\ 147 \\ 213 \\ 91 \\ 212 \\ 135 \\ 191 \\ 101 \\ 131 \\ 152 \\ 184 \\ 88 \\ 147 \\ 122 \\ 97 \end{pmatrix}$$

X: female? + male?

$$\begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix}$$



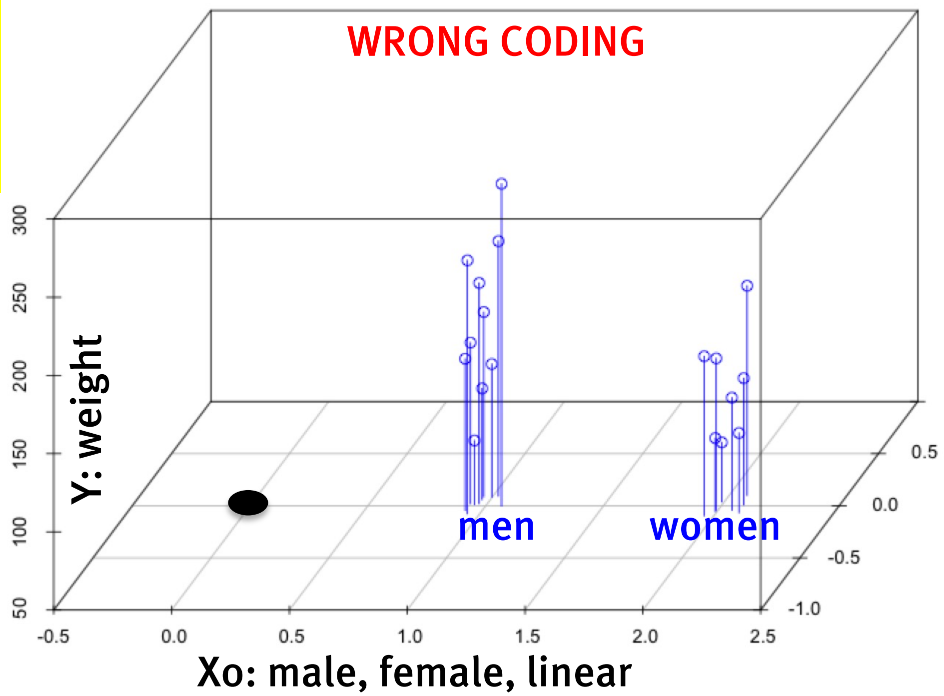
Geometric thinking about coefficients

	height	weight	sex
1	70	121	m
2	78	256	m
3	69	153	m
4	68	168	m
5	70	147	m
6	68	213	m
7	65	91	m
8	72	212	m
9	66	135	m
10	73	191	m
11	60	101	f
12	62	131	f
13	69	152	f
14	66	184	f
15	63	88	f
16	65	147	f
17	63	122	f
18	63	97	f

THIS IS WRONG!

Note that this means that
Mean(men) = 1*B1
Mean(women)=2*B1
Mean(women)-mean(men) = mean(men)

That's nonsense.



Y: weight

$$\begin{bmatrix} 121 \\ 256 \\ 153 \\ 168 \\ 147 \\ 213 \\ 91 \\ 212 \\ 135 \\ 191 \\ 101 \\ 131 \\ 152 \\ 184 \\ 88 \\ 147 \\ 122 \\ 97 \end{bmatrix}$$

X: male=1, female=2

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \end{bmatrix}$$

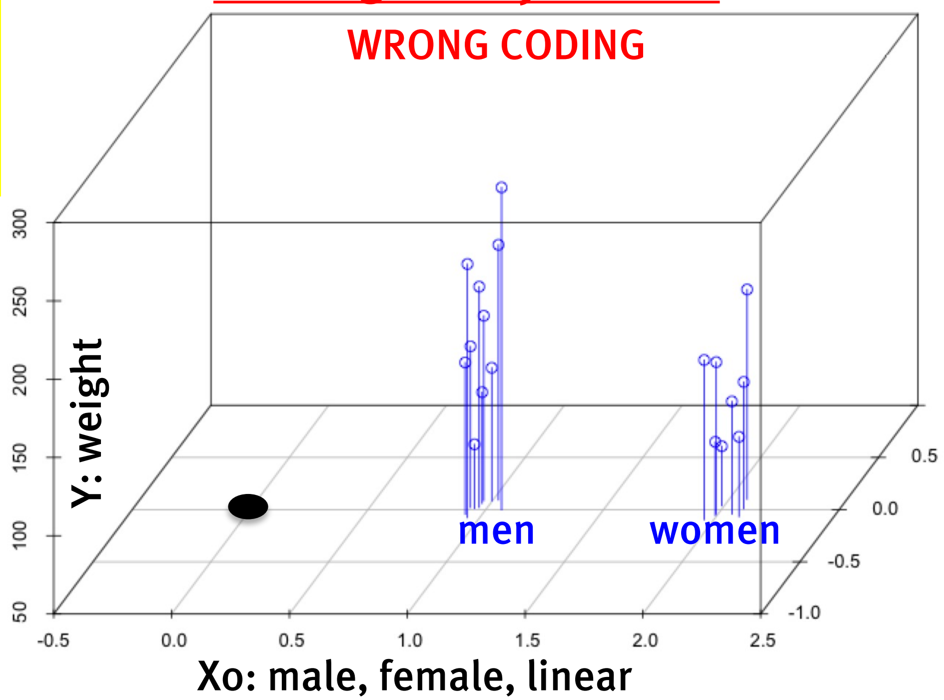
Geometric thinking about coefficients

	height	weight	sex
1	70	121	m
2	78	256	m
3	69	153	m
4	68	168	m
5	70	147	m
6	68	213	m
7	65	91	m
8	72	212	m
9	66	135	m
10	73	191	m
11	60	101	f
12	62	131	f
13	69	152	f
14	66	184	f
15	63	88	f
16	65	147	f
17	63	122	f
18	63	97	f

When coding categories with a number of regressors we need to be able to *independently* capture the difference between each category mean and 0 with the various coefficients.

If not, we get nonsense out.

Be careful when levels coded as integers in your data



Y: weight

$$\begin{pmatrix} 121 \\ 256 \\ 153 \\ 168 \\ 147 \\ 213 \\ 91 \\ 212 \\ 135 \\ 191 \\ 101 \\ 131 \\ 152 \\ 184 \\ 88 \\ 147 \\ 122 \\ 97 \end{pmatrix}$$

X: male=1, female=2

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \end{pmatrix}$$

R's default coding scheme

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix}$$

Intercept is the first factor level (default alphabetical order). Other coefficients are difference between nth level and the

```
sex
[18] m m m m m m m m m m f f f f f f f f
```

```
weight
[18] 121 256 153 168 147 213  91 212 135 191 101 131 152 184  88 147 122  97
```

```
summary(lm(weight~sex))
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   127.75      15.19   8.411 2.88e-07 ***
sexm           40.95      20.38   2.010   0.0617 .
```

The "m" indicates that this is coding for the offset of the "m" (here: male) category relative to the alphabetically first (here "f", female) category.

The estimate of the intercept is the estimated average female weight, and the estimate of the 'slope' or the 'sexm' coefficient is Mean(male)-Mean(female)

1-factor 2-levels: single-var regression

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix}$$

Intercept is the first (alphabetical) category.
Other coefficients are difference between nth category and the first one

```
summary(lm(weight~sex))
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   127.75      15.19   8.411 2.88e-07 ***
sexm           40.95      20.38   2.010   0.0617 .
```

This 'slope' is mean(males) minus mean(females). With a std. err. And a t-value. That's just a t-test. The same t-test we get if we assume equal var

```
t.test(weight~sex, var.equal=T)
```

```
	Two Sample t-test

data:  weight by sex
t = -2.0095, df = 16, p-value = 0.06166
```

F-statistic (comparing a model that codes for a gender difference to one that does not), is just the t-statistic squared. And the p-values are matched.

```
anova(lm(weight~sex))
```

```
Response: weight
          Df  Sum Sq Mean Sq F value  Pr(>F)
sex        1  7452.9  7452.9  4.0382 0.06166 .
Residuals 16 29529.6  1845.6
```

How does R code for categories?

```
      country height
1    North K.     62
2    North K.     73
3    North K.     64
4    Nort
5    Nort
6    Sout
7    Sout
8    Sout
9    Sout
10
11
12
13        USA
14        USA     70
15        USA     76
16 Netherlands    66
17 Netherlands    75
18 Netherlands    79
```

How would R code for country if you fit height~country?

```
summary(lm(height~country))
```

```
Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)        73.296      2.589  28.316 9.25e-14 ***
countryNorth K.    -5.849      3.274  -1.786   0.0957 .
countrySouth K.    -3.666      3.424  -1.070   0.3025
countryUSA         -4.057      3.170  -1.280   0.2214
```

Is that a hint?

What do the coefficients (and their significance) mean?

How does R code for categories?

```
        country height
1      North K.     62
2      North K.     73
3      North K.     64
4      North K.     67
5      North K.     71
6      South K.     72
7      South K.     71
8      South K.     72
9      South K.     64
10          USA     66
11          USA     66
12          USA     69
13          USA     68
14          USA     70
15          USA     76
16 Netherlands      66
17 Ne
18 Ne
```

```
(Intercept) countryNK countrySK countryUSA
          1         1         0          0
          1         1         0          0
          1         1         0          0
          1         1         0          0
          1         1         0          0
          1         0         1          0
          1         0         1          0
          1         0         1          0
          1         0         1          0
          1         0         0          1
          1         0         0          1
          1         0         0          1
          1         0         0          1
          1         0         0          1
          1         0         0          1
          1         0         0          0
```

```
summary(lm(height~country))

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)       73.296      2.589  28.316 9.25e-14 ***
countryNorth K.   -5.849      3.274  -1.786   0.0957 .
countrySouth K.   -3.666      3.424  -1.070   0.3025
countryUSA        -4.057      3.170  -1.280   0.2214
```


What do the coefficients mean?

How does R code for categories?

```
summary(lm(height~country))
```

```
Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)           73.296      2.589  28.316 9.25e-14 ***
countryNorth K.       -5.849      3.274  -1.786   0.0957 .
countrySouth K.       -3.666      3.424  -1.070   0.3025
countryUSA            -4.057      3.170  -1.280   0.2214
```

What do the coefficients mean?

Mean height of Netherlands is 73”

Mean height of N.K. is 5.8” shorter than Netherlands

Mean height of S.K. is 3.7” shorter than Netherlands.

Mean height of USA is 4” shorter than Netherlands

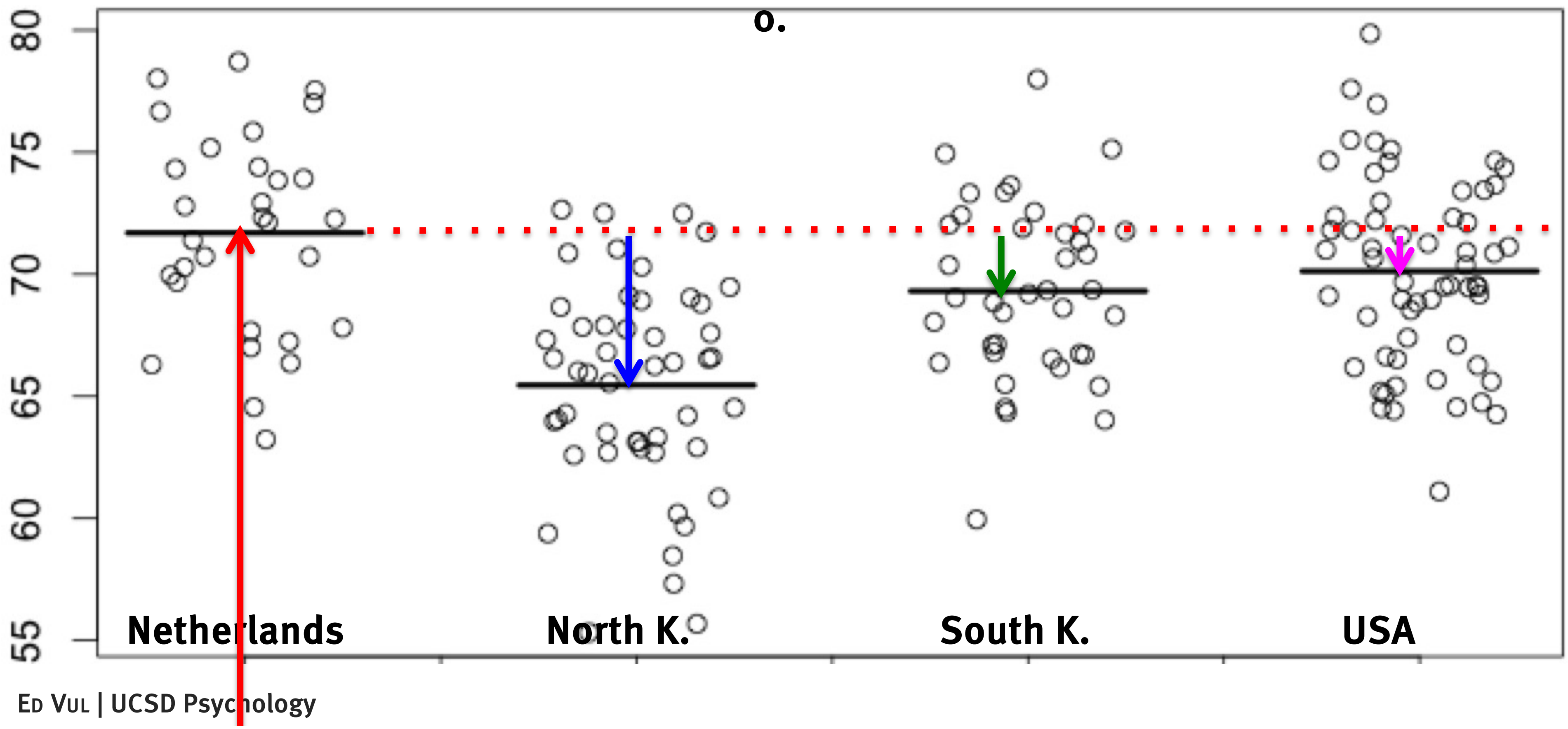
Mean height of Netherlands is significantly different from 0.

Differences between Netherlands and other countries are not significant.

Visualizing coefficients

```
summary(lm(height~country))

                 Estimate Std. Error t value Pr(>|t|)
(Intercept)      71.6960     0.7247  98.925  < 2e-16 ***
countryNorth K.  -6.2374     0.9167  -6.804 1.53e-10 ***
countrySouth K.  -2.3837     0.9588  -2.486   0.0138 *
countryUSA       -1.5696     0.8876  -1.768   0.0787 .
```

(Intercept): Mean height of Netherlands. Significance: comparison of Neth. mean to 0.



Categorical coefficient estimates

```
summary(lm(height~country))
Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)         73.296      2.589  28.316 9.25e-14 ***
countryNorth K.     -5.849      3.274  -1.786   0.0957 .
countrySouth K.     -3.666      3.424  -1.070   0.3025
countryUSA          -4.057      3.170  -1.280   0.2214
```

From this we learn:

Mean height of Netherlands is significantly different from 0.
Other pairwise differences with Netherlands are not significant.

But that's not what we want to know. We want to know:

Does mean height *vary* as a function of country?

So we do the F-test: An analysis of *variance* across means

Does the mean vary with a factor?

```
summary(lm(height~country))

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)        73.296      2.589  28.316 9.25e-14 ***
countryNorth K.    -5.849      3.274  -1.786   0.0957 .
countrySouth K.    -3.666      3.424  -1.070   0.3025
countryUSA         -4.057      3.170  -1.280   0.2214
```

But that's not what we want to know.
We want to know: does mean height vary as a function of country?

```
anova(lm(height~country))

Response: height
          Df  Sum Sq Mean Sq F value Pr(>F)
country    3  64.782  21.594  1.0743 0.3917
Residuals 14 281.414  20.101
```

It doesn't, but at least that's the answer we're after.

Does the mean vary with a factor?

```
anova(lm(height~country))
```

```
Response: height
          Df  Sum Sq Mean Sq F value Pr(>F)
country    3  64.782  21.594  1.0743 0.3917
Residuals 14 281.414  20.101
```

Note: df of country factor is not 1, but 3, because it takes 3 variables to code for differences among 4 categories.

F = SSR[country] / (4-1) / SSE[country] / (n-4)
p = 1-pf(F, 4-1, n-4)

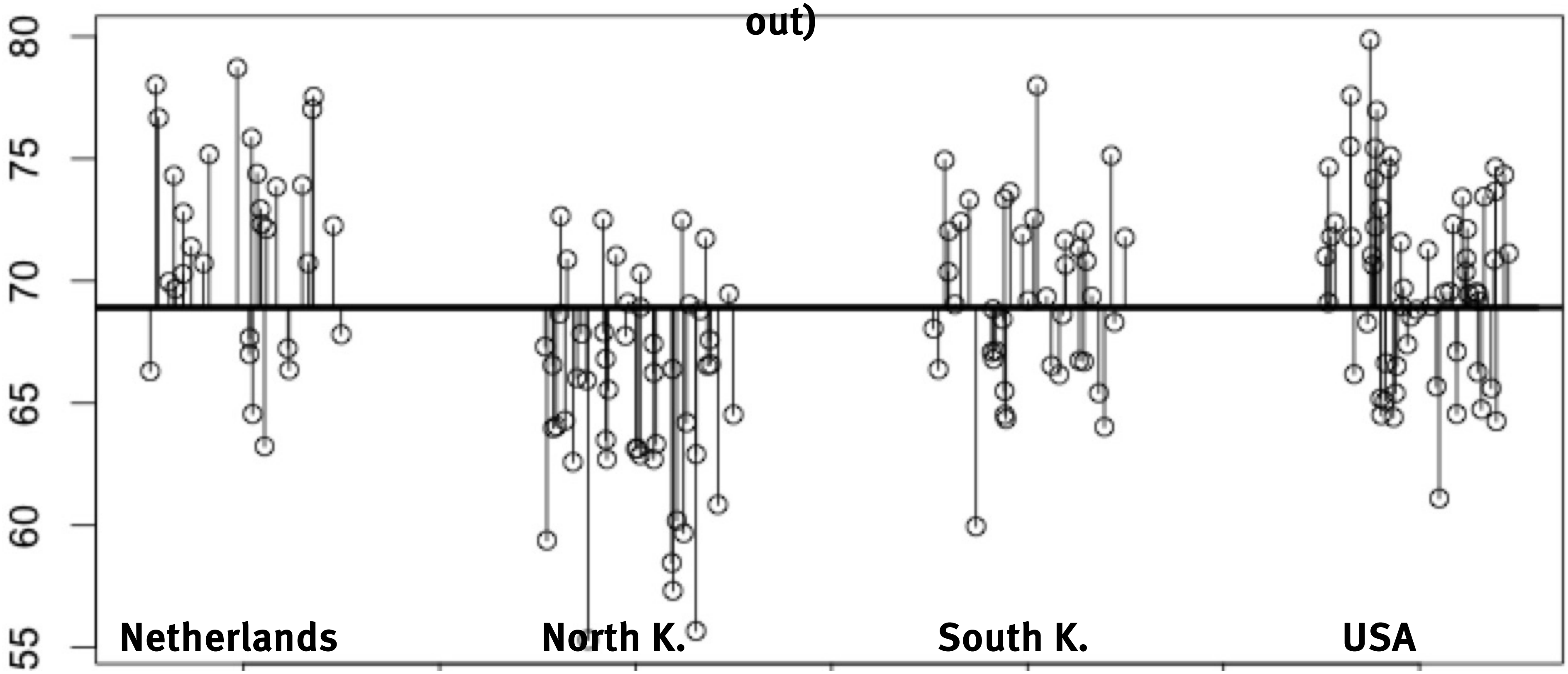
So, the country factor does not account for a significant amount of variance, compared to a model that only captures the average height.

Visualizing sums of squares

```
anova(lm(height~country))
```

```
Response: height
           Df  Sum Sq Mean Sq F value    Pr(>F)
country     3  923.72 307.906   19.54 5.567e-11 ***
Residuals 176 2773.38  15.758
```

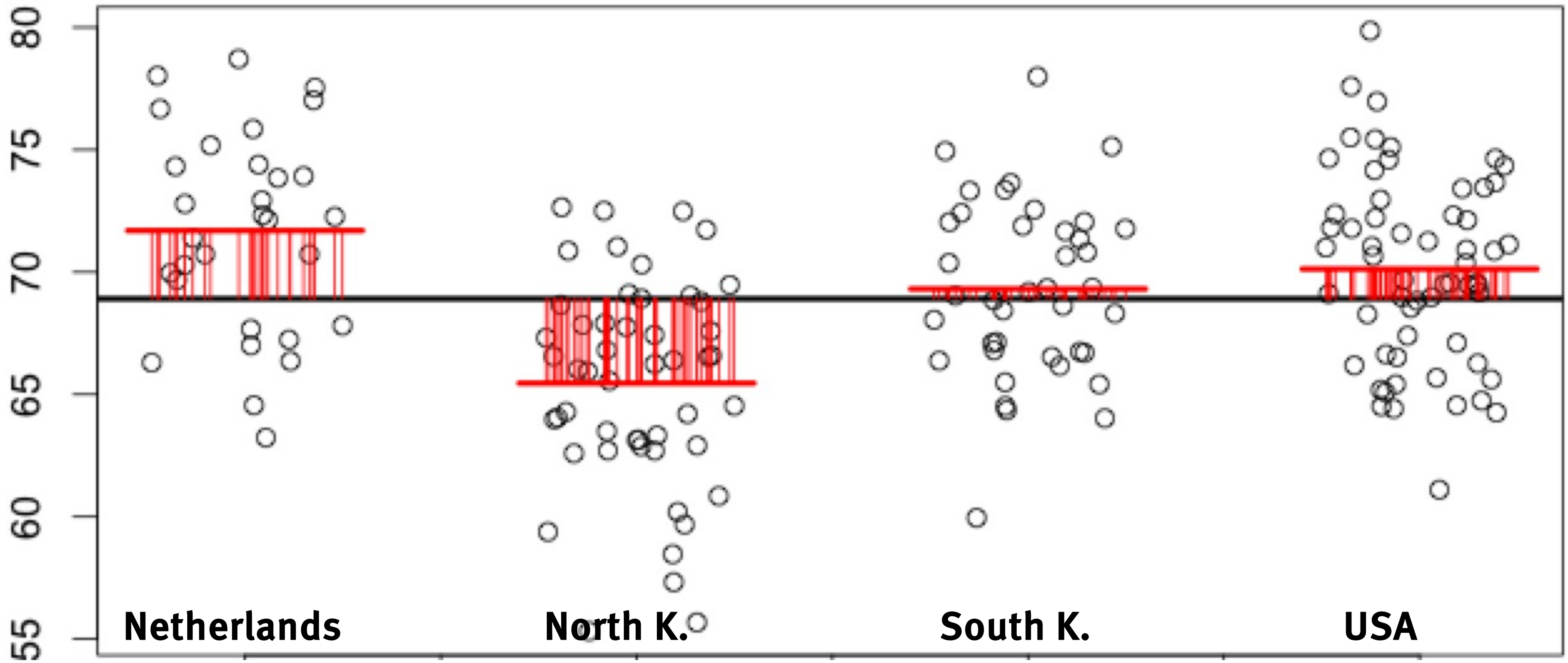
SST: sum of squared deviations of all data points from overall (grand) mean. (not in R out)



Visualizing sums of squares

```
anova(lm(height~country))
Response: height
           Df  Sum Sq Mean Sq F value    Pr(>F)
country     3  923.72 307.906   19.54 5.567e-11 ***
Residuals 176 2773.38  15.758
```

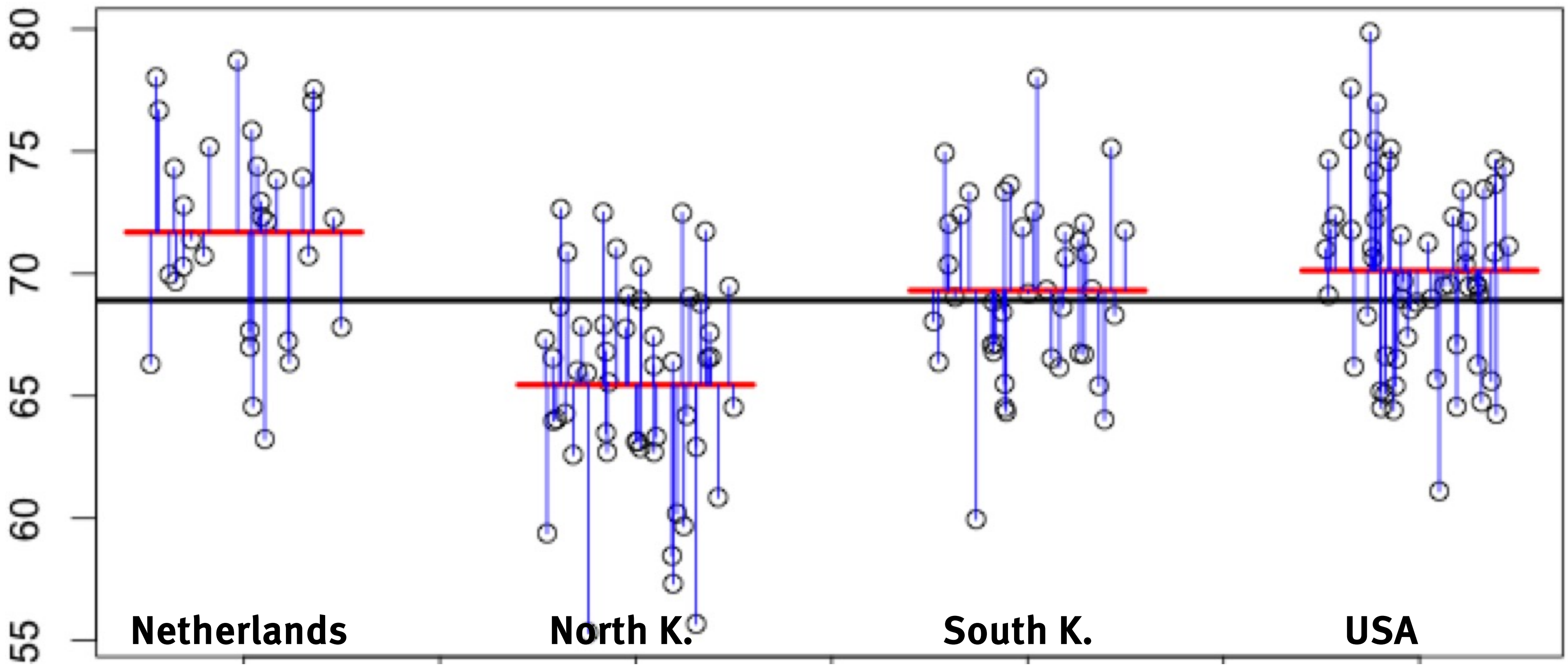
SSR[country]: sum(deviations^2) of country means from grand mean.
This is equivalent to Sum_country((mean(country) – grand_mean)^2*n_country)



Visualizing sums of squares

```
anova(lm(height~country))
Response: height
           Df  Sum Sq Mean Sq F value    Pr(>F)
country     3  923.72 307.906   19.54 5.567e-11 ***
Residuals 176 2773.38  15.758
```

SSE[country]: sum(deviations^2) of data points from respective country means.



Factor significance

```
anova(lm(height~country))
```

```
Response: height
           Df  Sum Sq Mean Sq F value    Pr(>F)
country     3  923.72 307.906   19.54 5.567e-11 ***
Residuals 176 2773.38  15.758
```

F test compares the SSR (or equivalently: SSE, or R^2) for a model that includes 3 regressors to capture country effects, to a null model where that SS allocation arises only from random variation due to residuals.

$$F(p_{SOURCE}, n - p_{FULL}) = \frac{\left(\frac{SSR_{SOURCE}}{p_{SOURCE}}\right)}{\left(\frac{SSE_{FULL}}{n - p_{FULL}}\right)}$$

```
F.Country = (923/3) / (2773/176)
```

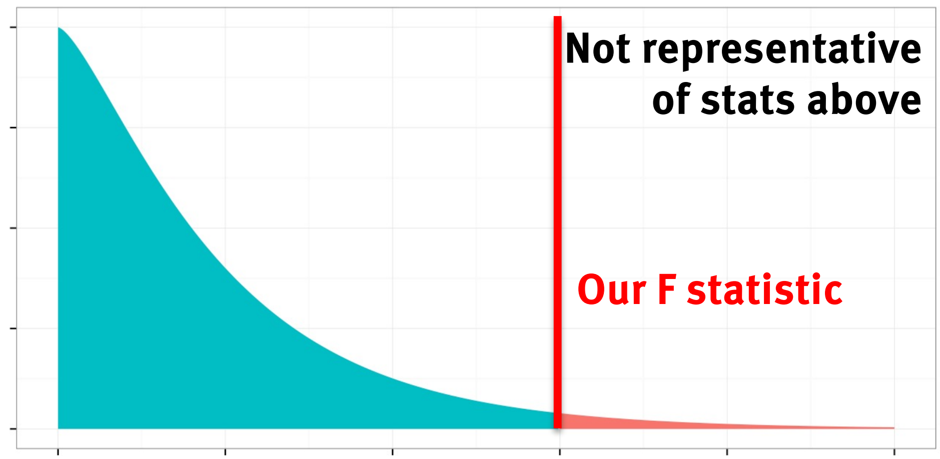
```
19.5
```

```
p.Country = 1-pf(19.54, 3, 176)
```

```
5e-11
```

F statistic measures how much variance is explained by factor.

More "signal variance" always means bigger F, so we do a one-tailed test.



Does the mean vary with a factor?

New data (n*10)

```
anova(lm(height~country))
```

```
Response: height
           Df  Sum Sq Mean Sq F value    Pr(>F)
country     3  923.72 307.906   19.54 5.567e-11 ***
Residuals 176 2773.38  15.758
```

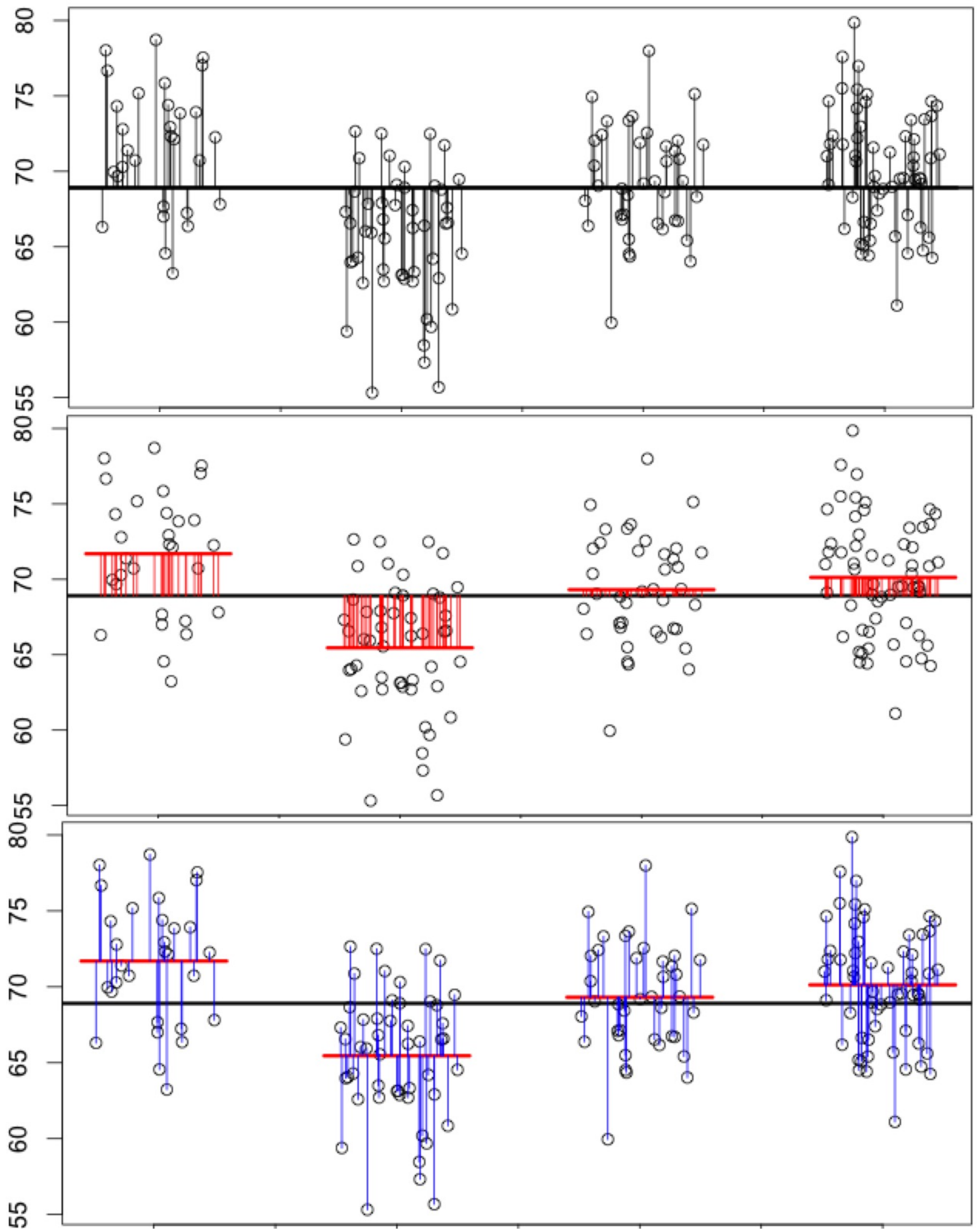
So now it's significant. What does that mean?

Equivalent statements:

(1) Variation of mean height among countries is significantly bigger than expected by chance if all means are really equal in population.

(2) Adding regressors to capture differences among countries accounts for more variance than expected by chance (because of 1!)

One way ANOVA summary.



As always:
SST = SSR + SSE
SSE = (1-R^2)*SST
R^2 = SSR/SST
although we now call it eta^2,
η^2

This is not just to mess with you – with more factors it ends up a bit different, but with one factor, it's the same.

As always with linear model, we calculate significance of SS allocation using the F statistic.

$$F(p_{SOURCE}, n - p_{FULL}) = \frac{\left(\frac{SSR_{SOURCE}}{p_{SOURCE}}\right)}{\left(\frac{SSE_{FULL}}{n - p_{FULL}}\right)}$$

```
summary(df)
```

```
 major          height
 cogs:10   Min.   :58.18
 ling:10   1st Qu.:62.62
 math:10   Median :65.08
 psyc:10   Mean   :65.09
 rady:10   3rd Qu.:67.55
           Max.   :71.73
```

```
anova(lm(data=df, height~major))
```

```
Response: height
          Df Sum Sq
major      4 397.04
Residuals 45 786.75
```

```
summary(lm(data=df, height~major))
```

```
Coefficients:
            Estimate Std. Error
(Intercept)  69.6589     1.3222
majorling    -1.5687     1.8699
majormath    -7.4371     1.8699
majorpsyc     0.4074     1.8699
majorrady    -2.7078     1.8699
```

- What's the mean height of cogs majors?
- What's the mean height of math majors?
- What's the difference between mean height of psyc and rady?
- What's the t-test statistic and significance of the "math" coefficient? What does it mean?
- What's effect size (eta^2 / R^2) of major on height?
- Is the ANOVA on the major factor significant? What's the F statistic? P-value?

```
t.test(df$height[df$major=='math'], df$height[df$major=='cogs'])
```

```
t = -3.8896, df = 17.922, p-value = 0.001081
```

```
t.test(df$height[df$major=='math'], df$height[df$major=='cogs'], var.equal = T)
```

```
t = -3.8896, df = 18, p-value = 0.001074
```

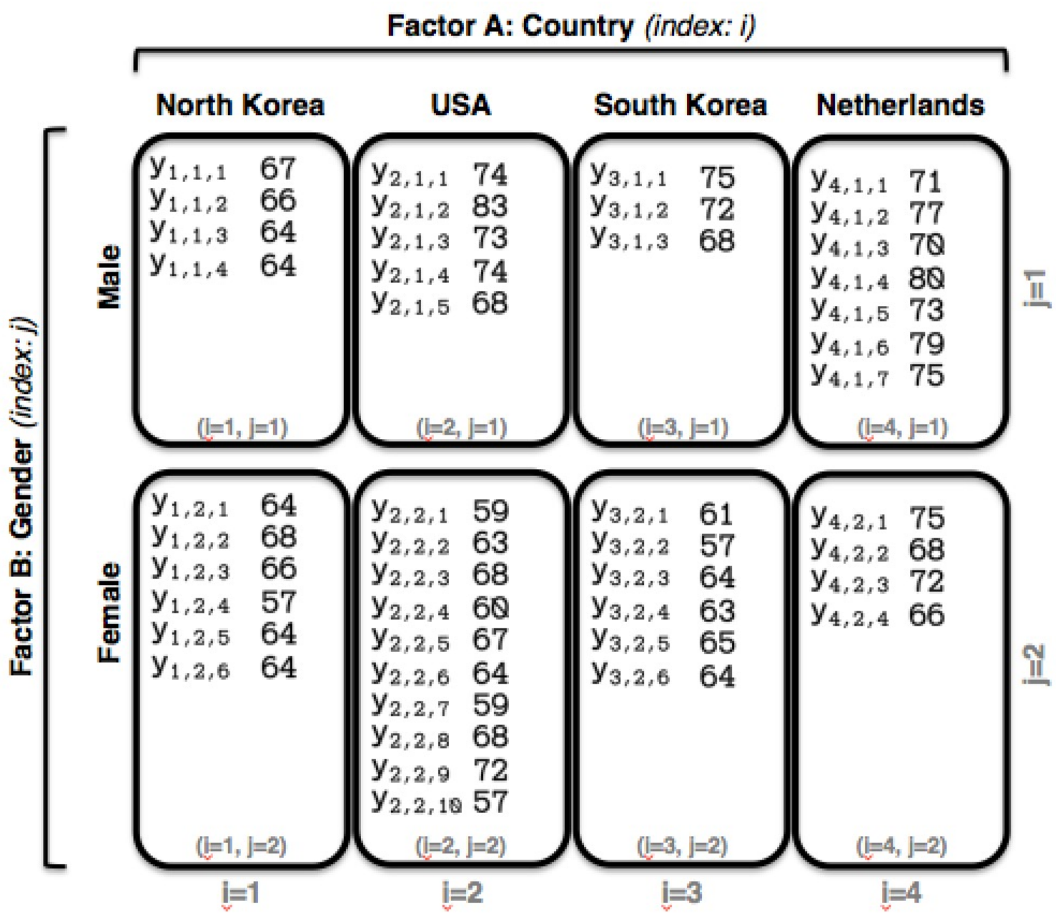
```
summary(lm(data=df, height~major))
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  69.6589     1.3222  52.682  < 2e-16 ***
majorling    -1.5687     1.8699  -0.839  0.40597
majormath    -7.4371     1.8699  -3.977  0.00025 ***
majorpsyc     0.4074     1.8699   0.218  0.82850
majorrady    -2.7078     1.8699  -1.448  0.15453
```

- What's the difference between the eq. var t-test of math-cogs and the t-test on the math coefficient?

Representing factorial designs

	height	sex	country
1	62	f	N.Korea
2	57	f	N.Korea
3	60	f	N.Korea
4	57	f	N.Korea
5	59	f	N.Korea
6	67	m	S.Korea
7	61	m	S.Korea
8	57	m	S.Korea
9	68	m	S.Korea
10	60	f	USA
11	60	f	USA
12	60	f	USA
13	64	f	USA
14	65	f	USA
15	74	m	Netherlands
16	69	m	Netherlands
17	62	m	Netherlands
18	74	m	Netherlands
19	63	m	Netherlands
20	59	f	N.Korea
21	63	f	N.Korea
22	67	f	N.Korea
23	68	f	N.Korea
24	72	f	N.Korea
25	61	f	N.Korea
26	63	m	S.Korea
27	72	m	S.Korea
28	67	m	S.Korea
29	67	m	S.Korea
30	64	f	USA
31	64	f	USA
32	65	f	USA
33	63	f	USA
34	56	f	USA
35	64	f	USA
36	68	m	Netherlands
37	67	m	Netherlands
38	72	m	Netherlands
39	71	m	Netherlands
40	73	m	Netherlands
41	74	m	Netherlands



(Intercept)	c:Neth	c:S.K.	c:USA	sex:m
1	0	0	0	0
1	0	0	0	0
1	0	0	0	0
1	0	0	0	0
1	0	0	0	0
1	0	0	0	1
1	0	0	0	1
1	0	0	0	1
1	0	0	0	1
1	0	1	0	0
1	0	1	0	0
1	0	1	0	0
1	0	1	0	0
1	0	1	0	0
1	0	1	0	1
1	0	1	0	1
1	0	1	0	1
1	0	1	0	1
1	0	1	0	1
1	0	0	1	0
1	0	0	1	0
1	0	0	1	0
1	0	0	1	0
1	0	0	1	0
1	0	0	1	0
1	0	0	1	1
1	0	0	1	1
1	0	0	1	1
1	0	0	1	1
1	1	0	0	0
1	1	0	0	0
1	1	0	0	0
1	1	0	0	0
1	1	0	0	0
1	1	0	0	0
1	1	0	0	1
1	1	0	0	1
1	1	0	0	1
1	1	0	0	1
1	1	0	0	1
1	1	0	0	1

(Intercept)	c:Neth	c:S.K.	c:USA	sex:m
1	0	0	0	0
1	0	0	0	0
1	0	0	0	0
1	0	0	0	0
1	0	0	0	0
1	0	0	0	1
1	0	0	0	1
1	0	0	0	1
1	0	0	0	1
1	0	1	0	0
1	0	1	0	0
1	0	1	0	0
1	0	1	0	0
1	0	1	0	0
1	0	1	0	1
1	0	1	0	1
1	0	1	0	1
1	0	1	0	1
1	0	1	0	1
1	0	0	1	0
1	0	0	1	0
1	0	0	1	0
1	0	0	1	0
1	0	0	1	0
1	0	0	1	0
1	0	0	1	1
1	0	0	1	1
1	0	0	1	1
1	0	0	1	1
1	1	0	0	0
1	1	0	0	0
1	1	0	0	0
1	1	0	0	0
1	1	0	0	0
1	1	0	0	0
1	1	0	0	1
1	1	0	0	1
1	1	0	0	1
1	1	0	0	1
1	1	0	0	1
1	1	0	0	1

<- Coding just for "main effects": additive effects of a factor.
Main effect of sex: average difference between men and women
Main effect of country: average differences between countries.

```
summary(lm(height~country+sex))

 Estimate Std. Error t value Pr(>|t|)
(Intercept)          58.437       1.429   40.891   < 2e-16 ***
countryNetherlands    5.555       1.745    3.183   0.00300 **
countryS.Korea        3.905       1.818    2.148   0.03855 *
countryUSA            5.256       1.818    2.892   0.00646 **
sexm                  5.517       1.243    4.439  8.22e-05 ***
```

So, the model predicts different cell means to be:

```
N.K. females = B0                         (intercept)
Netherlands females = B0 + B1             + (countryNetherlands)
S.K. females = B0 + B2                    + (countryS.Korea)
USA females = B0 + B3                     + (countryUSA)
N.K. males = B0 + B4                      + (sexm)
Netherlands males = B0 + B1 + B4           + (netherlands) + (sexm)
S.K. males = B0 + B2 + B4                 + (S.K.) + (sexm)
USA males = B0 + B3 + B4                  + (USA) + (sexm)
```

"main effects":

Effect of maleness is additive with effect of country.

Difference between males and females is the same for every country, and differences among countries are the same within males and within females.

(Intercept)	c:Neth	c:S.K.	c:USA	sex:m
1	0	0	0	0
1	0	0	0	0
1	0	0	0	0
1	0	0	0	0
1	0	0	0	0
1	0	0	0	1
1	0	0	0	1
1	0	0	0	1
1	0	0	0	1
1	0	1	0	0
1	0	1	0	0
1	0	1	0	0
1	0	1	0	0
1	0	1	0	0
1	0	1	0	1
1	0	1	0	1
1	0	1	0	1
1	0	1	0	1
1	0	1	0	1
1	0	0	1	0
1	0	0	1	0
1	0	0	1	0
1	0	0	1	0
1	0	0	1	0
1	0	0	1	0
1	0	0	1	1
1	0	0	1	1
1	0	0	1	1
1	0	0	1	1
1	1	0	0	0
1	1	0	0	0
1	1	0	0	0
1	1	0	0	0
1	1	0	0	0
1	1	0	0	0
1	1	0	0	1
1	1	0	0	1
1	1	0	0	1
1	1	0	0	1
1	1	0	0	1
1	1	0	0	1

<- Coding just for “main effects”: additive effects of a factor.
Main effect of sex: average difference between men and women
Main effect of country: average differences between countries.

```
summary(lm(height~country+sex))
```

```
 Estimate Std. Error t value Pr(>|t|)
(Intercept)          58.437      1.429  40.891  < 2e-16 ***
countryNetherlands    5.555      1.745   3.183  0.00300 **
countryS.Korea        3.905      1.818   2.148  0.03855 *
countryUSA            5.256      1.818   2.892  0.00646 **
sexm                  5.517      1.243   4.439 8.22e-05 ***
```

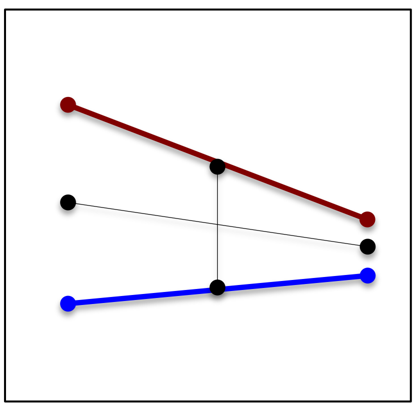
```
anova(lm(height~country+sex))
```

```
Response: height
          Df Sum Sq Mean Sq F value    Pr(>F)
country    3 196.18  65.394  4.1827   0.01223 *
sex        1 308.09 308.095 19.7060 8.217e-05 ***
Residuals 36 562.84  15.635
```

Significance of main effects (in ANOVA) says variation in average height across country is significantly greater than 0. Similarly, variation in average height across sex is greater than 0.

What does a sig. main effect mean?

1. Amount of variance accounted for by factor levels is bigger than chance.
2. Variance of means across factor level is greater than zero.
3. Evidence that not all factor level means are equal.



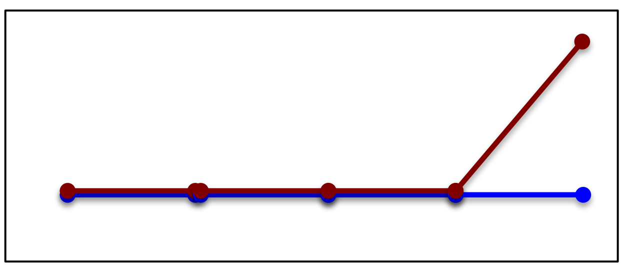
Compare mean of left vs right, and mean of red vs blue…

What does a sig. main effect mean?

1. Amount of variance accounted for by factor levels is bigger than chance.
2. Variance of means across factor level is greater than zero.
3. Evidence that not all factor level means are equal.

What it does not mean:

- That there is a uniform additive offset of factor level. (just one rogue cell would do)
- Or that the means vary in any other particular pattern. (mean changes might not coincide with your prediction)



Ugh: main effects will show up, but they aren't consistent with intuitive interpretation.

(Intercept)	c:Neth	c:S.K.	c:USA	sex:m
1	0	0	0	0
1	0	0	0	0
1	0	0	0	0
1	0	0	0	0
1	0	0	0	0
1	0	0	0	1
1	0	0	0	1
1	0	0	0	1
1	0	0	0	1
1	0	1	0	0
1	0	1	0	0
1	0	1	0	0
1	0	1	0	0
1	0	1	0	0
1	0	1	0	1
1	0	1	0	1
1	0	1	0	1
1	0	1	0	1
1	0	1	0	1
1	0	0	1	0
1	0	0	1	0
1	0	0	1	0
1	0	0	1	0
1	0	0	1	0
1	0	0	1	0
1	0	0	1	1
1	0	0	1	1
1	0	0	1	1
1	0	0	1	1
1	1	0	0	0
1	1	0	0	0
1	1	0	0	0
1	1	0	0	0
1	1	0	0	0
1	1	0	0	0
1	1	0	0	1
1	1	0	0	1
1	1	0	0	1
1	1	0	0	1
1	1	0	0	1
1	1	0	0	1

```
anova(lm(height~country+sex))
```

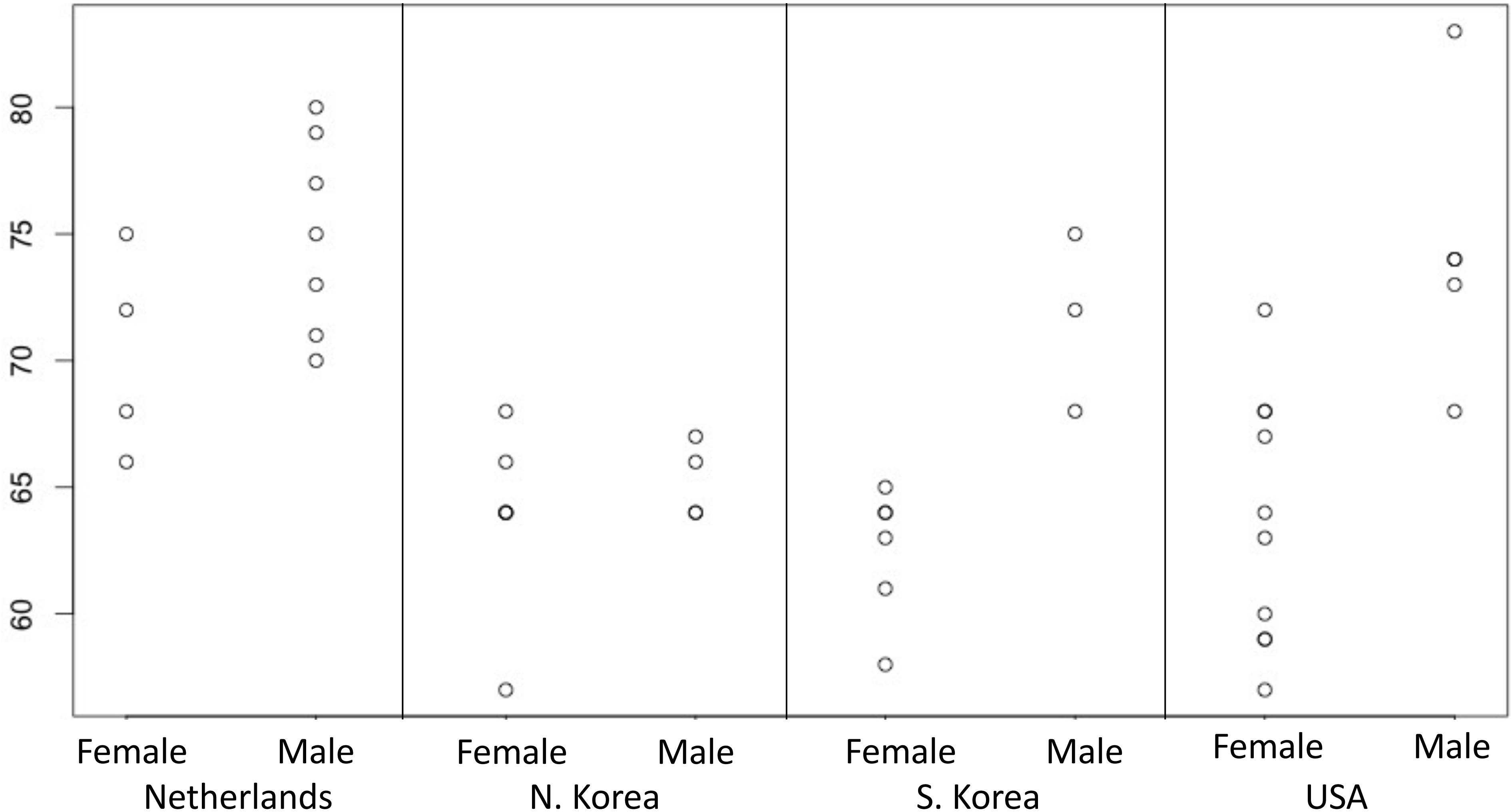
```
Response: height
          Df Sum Sq Mean Sq F value    Pr(>F)
country    3 196.18  65.394  4.1827   0.01223 *
sex        1 308.09 308.095 19.7060 8.217e-05 ***
Residuals 36 562.84  15.635
```

"Main effects"

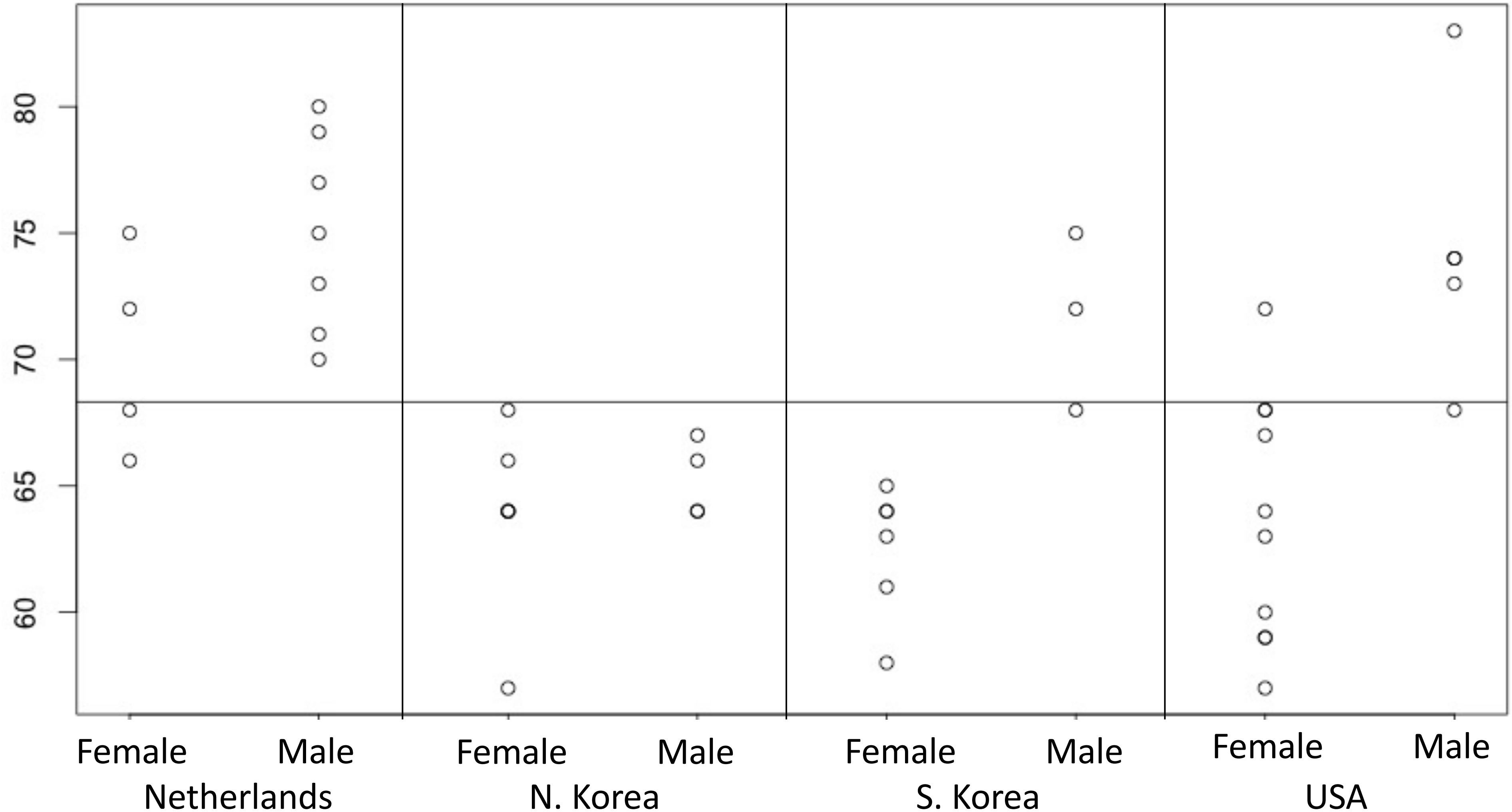
Effect of maleness is additive with effect of country.

Difference between males and females is the same for every country, and differences among countries are the same within males and within females.

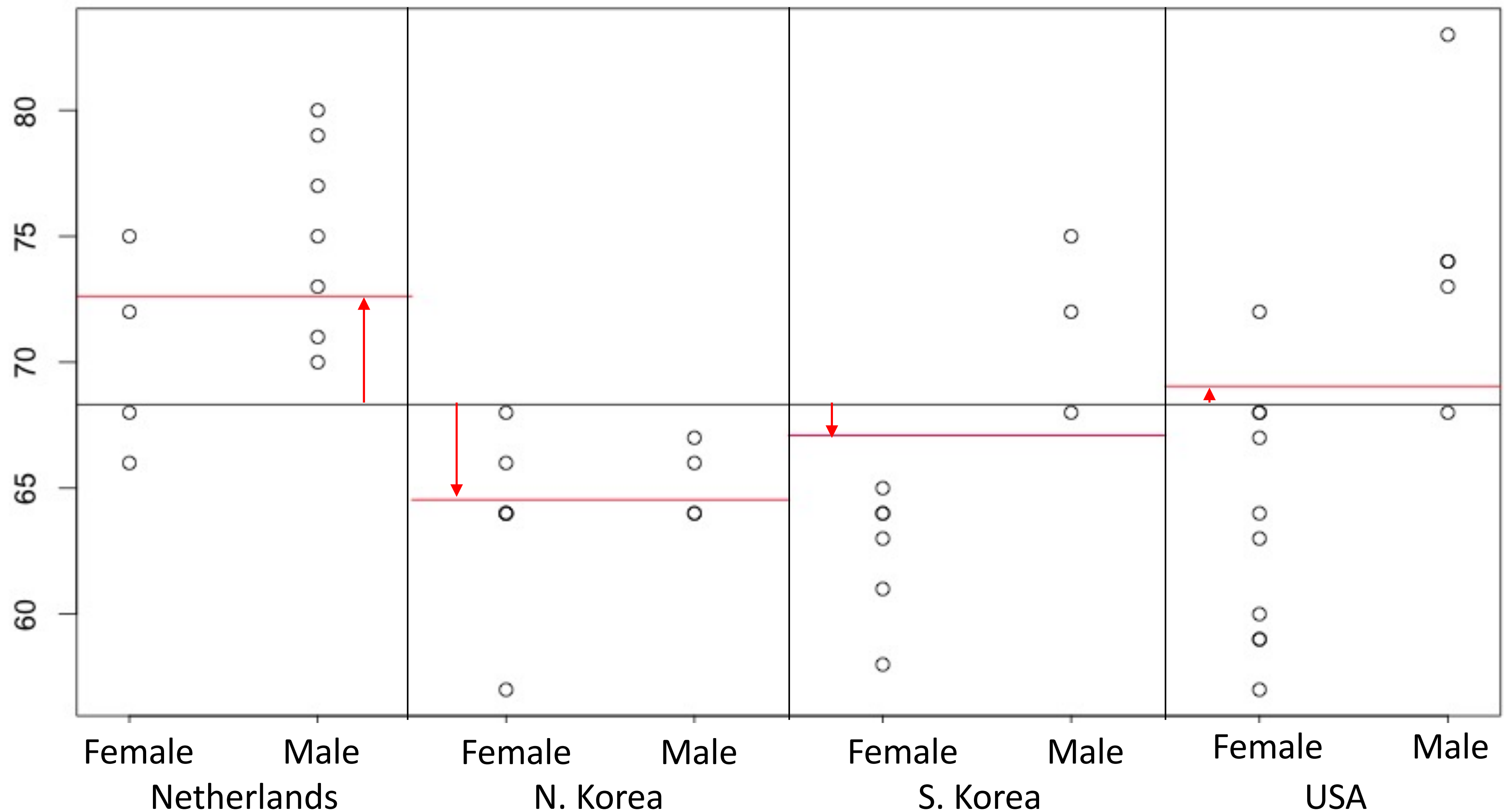
But, critically, this cannot capture "interactions" some differences in differences among means. E.g., mean(male)-mean(female) varies across countries.



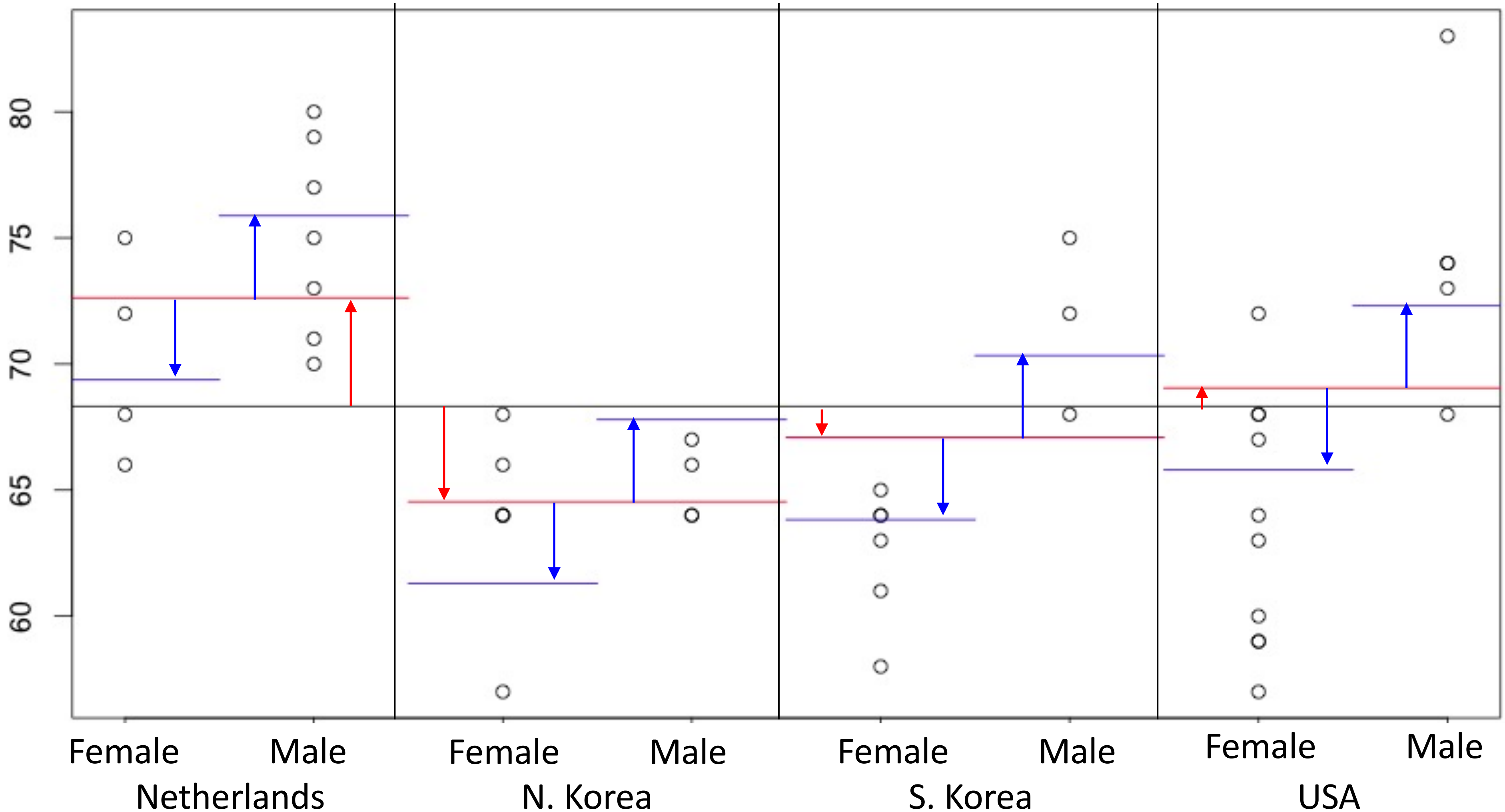
All the data (smaller design)



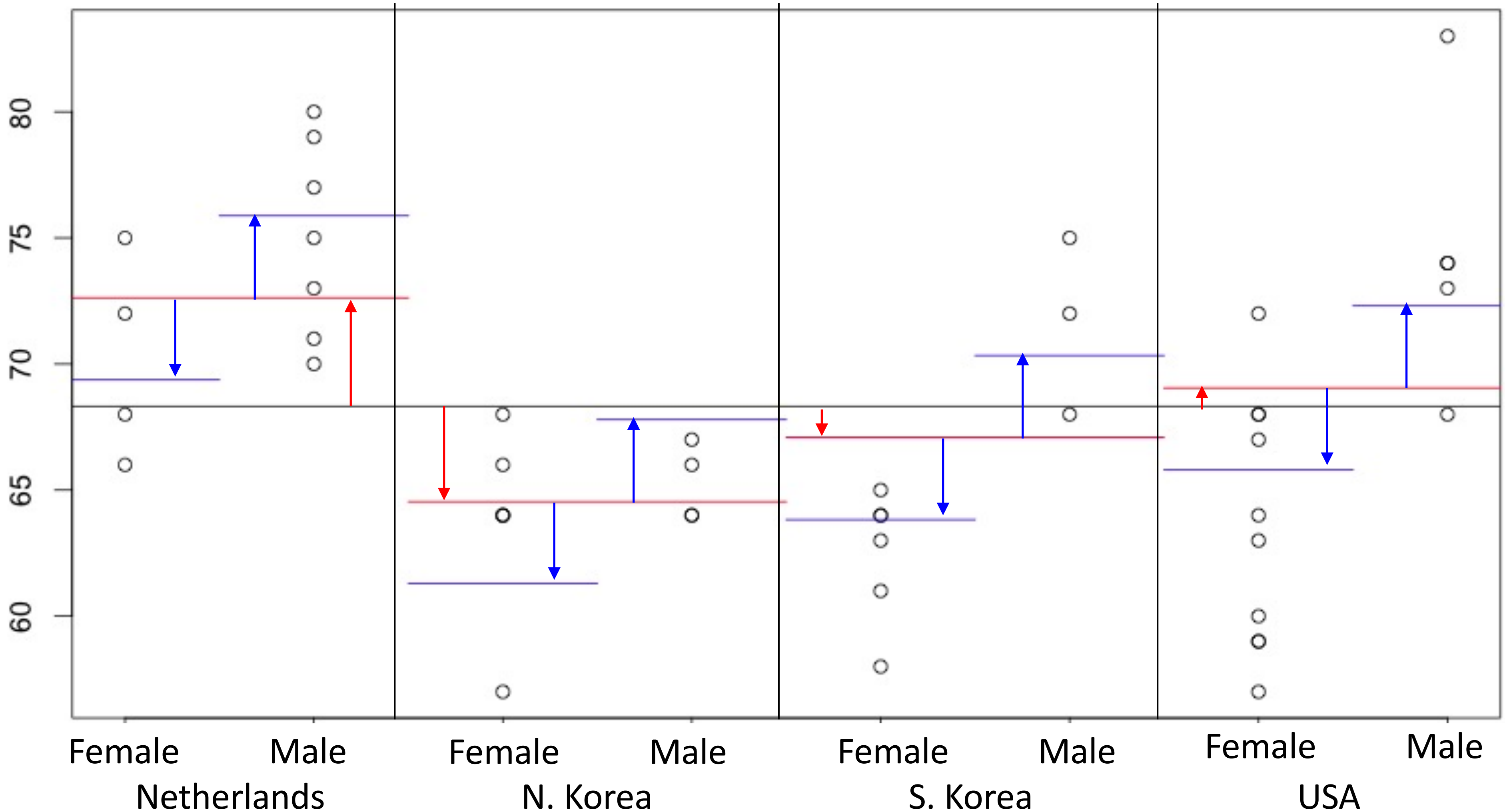
The overall mean.



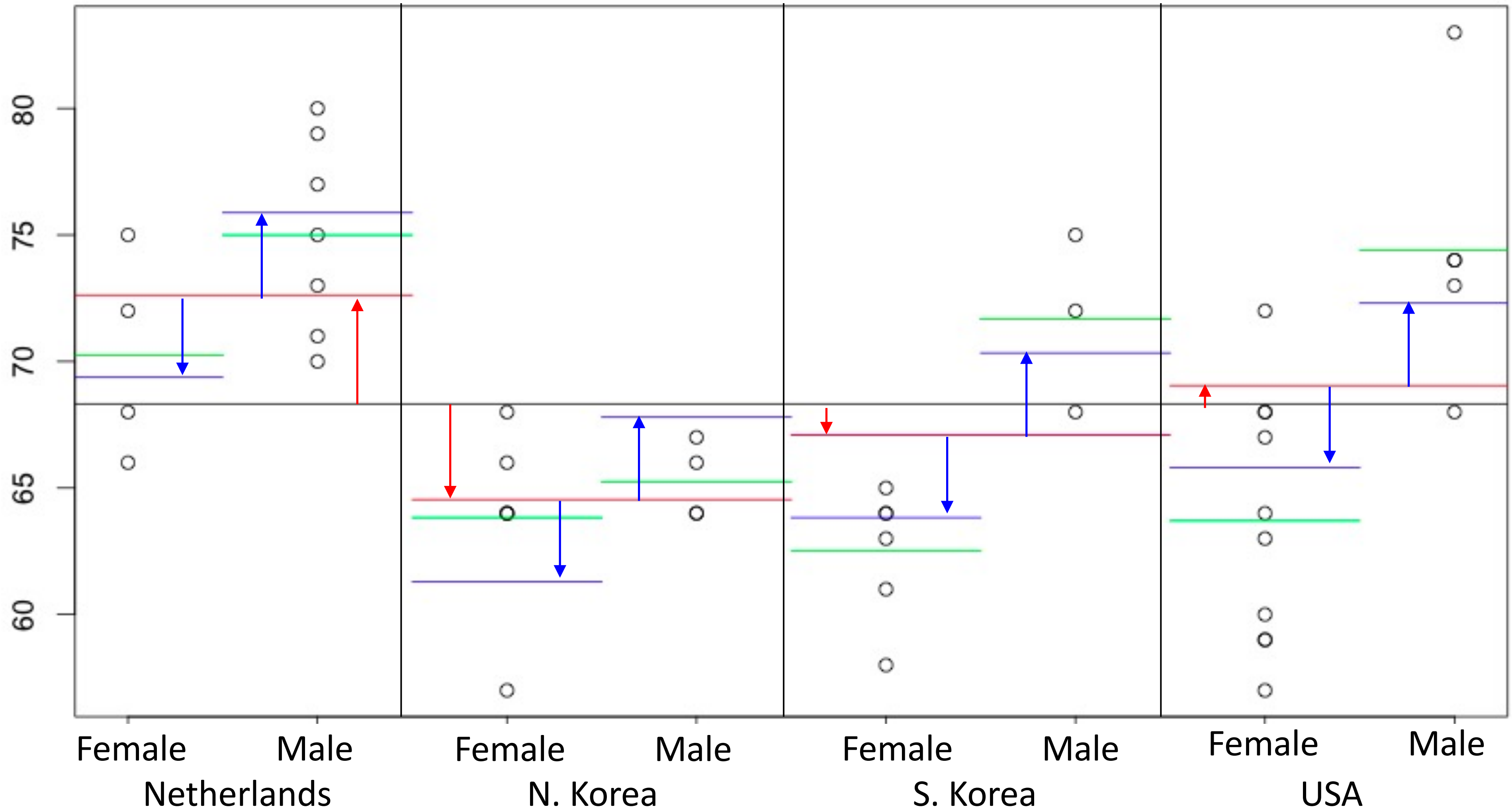
Main effects capture deviations of specific factor level means from overall mean



Main effects capture deviations of specific factor level means from overall mean



So the treatment 'main effects' are additive offsets for each treatment 'level' that are constant for all conditions at that treatment level.



So the treatment 'main effects' are offsets for each treatment 'level' that are constant for all conditions at that treatment level and additive across factors.

But they don't necessarily match the cell means. The distance left over is the "interaction".

```
anova(lm(height~country+sex+country:sex))
```

The same regressors we had before, coding for main effects

New regressors added to capture “interaction”

(Intercept)	c.Neth	c.S.K.	c.USA	sexM	c.Neth:sexM	c.S.K.:sexM	c.USA:sexM
1	0	0	0	0	0	0	[illegible]
1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
1	0	0	0	1	0	0	0
1	0	0	0	1	0	0	0
1	0	0	0	1	0	0	0
1	0	0	0	1	0	0	0
1	0	1	0	0	0	0	0
1	0	1	0	0	0	0	0
1	0	1	0	0	0	0	0
1	0	1	0	0	0	0	0
1	0	1	0	0	0	0	0
1	0	1	0	1	0	1	0
1	0	1	0	1	0	1	0
1	0	1	0	1	0	1	0
1	0	1	0	1	0	1	0
1	0	1	0	1	0	1	0
1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0
1	0	0	1	1	0	0	1
1	0	0	1	1	0	0	1
1	0	0	1	1	0	0	1
1	0	0	1	1	0	0	1
1	1	0	0	0	0	0	0
1	1	0	0	0	0	0	0
1	1	0	0	0	0	0	0
1	1	0	0	0	0	0	0
1	1	0	0	0	0	0	0
1	1	0	0	0	0	0	0
1	1	0	0	1	1	0	0
1	1	0	0	1	1	0	0
1	1	0	0	1	1	0	0
1	1	0	0	1	1	0	0
1	1	0	0	1	1	0	0
1	1	0	0	1	1	0	0

Adding A:B to the linear model adds the necessary indicator variables to capture the interaction.

- Different indicator variable designs can capture the interaction (yielding different coefficient interpretations)
- All capture unique mean in each cell.
- It takes (a-1)*(b-1) indicators to capture an interaction (where a = # levels in factor A)
- The full interaction model, we will have a*b regressors (including intercept): one for each cell.

(Intercept)	c.Neth	c.S.K.	c.USA	sexM	c.Neth:sexM	c.S.K.:sexM	c.USA:sexM
1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	
1	0	0	0	0	0	0	
1	0	0	0	0	0	0	
1	0	0	0	0	0	0	
1	0	0	0	1	0	0	
1	0	0	0	1	0	0	
1	0	0	0	1	0	0	
1	0	0	0	1	0	0	
1	0	1	0	0	0	0	
1	0	1	0	0	0	0	
1	0	1	0	0	0	0	
1	0	1	0	0	0	0	
1	0	1	0	0	0	0	0
1	0	1	0	1	0	1	0
1	0	1	0	1	0	1	0
1	0	1	0	1	0	1	0
1	0	1	0	1	0	1	0
1	0	1	0	1	0	1	0
1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0
1	0	0	1	1	0	0	1
1	0	0	1	1	0	0	1
1	0	0	1	1	0	0	1
1	0	0	1	1	0	0	1
1	1	0	0	0	0	0	0
1	1	0	0	0	0	0	0
1	1	0	0	0	0	0	0
1	1	0	0	0	0	0	0
1	1	0	0	0	0	0	0
1	1	0	0	0	0	0	0
1	1	0	0	1	1	0	0
1	1	0	0	1	1	0	0
1	1	0	0	1	1	0	0
1	1	0	0	1	1	0	0
1	1	0	0	1	1	0	0
1	1	0	0	1	1	0	0

```
anova(lm(height~country+sex+country:sex))
```

```
Response: height
            Df Sum Sq Mean Sq F value    Pr(>F)
country      3 196.18  65.394  4.2342   0.01226 *
sex          1 308.09 308.095 19.9486 8.803e-05 ***
country:sex  3  53.18  17.726  1.1477   0.34436
Residuals   33 509.67  15.444
```

So, here we have Type I sums of squares results

The interpretation is:

- Adding country regressors to a null (grand mean) model accounts for significantly more variation than expected by chance. (variation in mean height across countries is greater than 0)
- Adding sex regressors to a model with country accounts for significantly more variation (variation in mean height across sex is greater than 0)
- Adding country:sex interaction regressors to a model with country and sex main effects does not account for significantly more variation (pattern of mean differences across countries is not significantly different for males than females)

(Intercept)	c.Neth	c.S.K.	c.USA	sexM	c.Neth:sexM	c.S.K.:sexM	c.USA:sexM
1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	
1	0	0	0	0	0	0	
1	0	0	0	0	0	0	
1	0	0	0	0	0	0	
1	0	0	0	1	0	0	
1	0	0	0	1	0	0	
1	0	0	0	1	0	0	
1	0	0	0	1	0	0	
1	0	1	0	0	0	0	
1	0	1	0	0	0	0	
1	0	1	0	0	0	0	
1	0	1	0	0	0	0	
1	0	1	0	0	0	0	0
1	0	1	0	1	0	1	0
1	0	1	0	1	0	1	0
1	0	1	0	1	0	1	0
1	0	1	0	1	0	1	0
1	0	1	0	1	0	1	0
1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0
1	0	0	1	1	0	0	1
1	0	0	1	1	0	0	1
1	0	0	1	1	0	0	1
1	0	0	1	1	0	0	1
1	1	0	0	0	0	0	
1	1	0	0	0	0	0	
1	1	0	0	0	0	0	
1	1	0	0	0	0	0	
1	1	0	0	0	0	0	
1	1	0	0	0	0	0	
1	1	0	0	1	1	0	
1	1	0	0	1	1	0	
1	1	0	0	1	1	0	
1	1	0	0	1	1	0	
1	1	0	0	1	1	0	
1	1	0	0	1	1	0	

```
anova(lm(height~country+sex+country:sex))
```

```
Response: height
            Df Sum Sq Mean Sq F value    Pr(>F)    
country      3 196.18  65.394  4.2342   0.01226 *  
sex          1 308.09 308.095 19.9486 8.803e-05 ***
country:sex  3  53.18  17.726  1.1477   0.34436    
Residuals   33 509.67  15.444
```

We can adopt a shortcut in R to get the full model

```
anova(lm(height~country*sex))
```

```
Response: height
            Df Sum Sq Mean Sq F value    Pr(>F)    
country      3 196.18  65.394  4.2342   0.01226 *  
sex          1 308.09 308.095 19.9486 8.803e-05 ***
country:sex  3  53.18  17.726  1.1477   0.34436    
Residuals   33 509.67  15.444
```

(Intercept)	c.Neth	c.S.K.	c.USA	sexM	c.Neth:sexM	c.S.K.:sexM	c.USA:sexM
1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	
1	0	0	0	0	0	0	
1	0	0	0	0	0	0	
1	0	0	0	0	0	0	
1	0	0	0	1	0	0	
1	0	0	0	1	0	0	
1	0	0	0	1	0	0	
1	0	0	0	1	0	0	
1	0	1	0	0	0	0	
1	0	1	0	0	0	0	
1	0	1	0	0	0	0	
1	0	1	0	0	0	0	
1	0	1	0	0	0	0	0
1	0	1	0	1	0	1	0
1	0	1	0	1	0	1	0
1	0	1	0	1	0	1	0
1	0	1	0	1	0	1	0
1	0	1	0	1	0	1	0
1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0
1	0	0	1	0	0	0	0
1	0	0	1	1	0	0	1
1	0	0	1	1	0	0	1
1	0	0	1	1	0	0	1
1	0	0	1	1	0	0	1
1	1	0	0	0	0	0	0
1	1	0	0	0	0	0	0
1	1	0	0	0	0	0	0
1	1	0	0	0			
1	1	0	0	0			
1	1	0	0	0			
1	1	0	0	1			
1	1	0	0	1			
1	1	0	0	1			
1	1	0	0	1			
1	1	0	0	1			
1	1	0	0	1			

```
anova(lm(height~country+sex+country:sex))
```

```
Response: height
             Df Sum Sq Mean Sq F value    Pr(>F)
country       3 196.18  65.394  4.2342   0.01226 *
sex           1 308.09 308.095 19.9486 8.803e-05 ***
country:sex   3  53.18  17.726  1.1477   0.34436
Residuals    33 509.67  15.444
```

Interpreting coefficients with interactions is weird and depends on how they are coded.

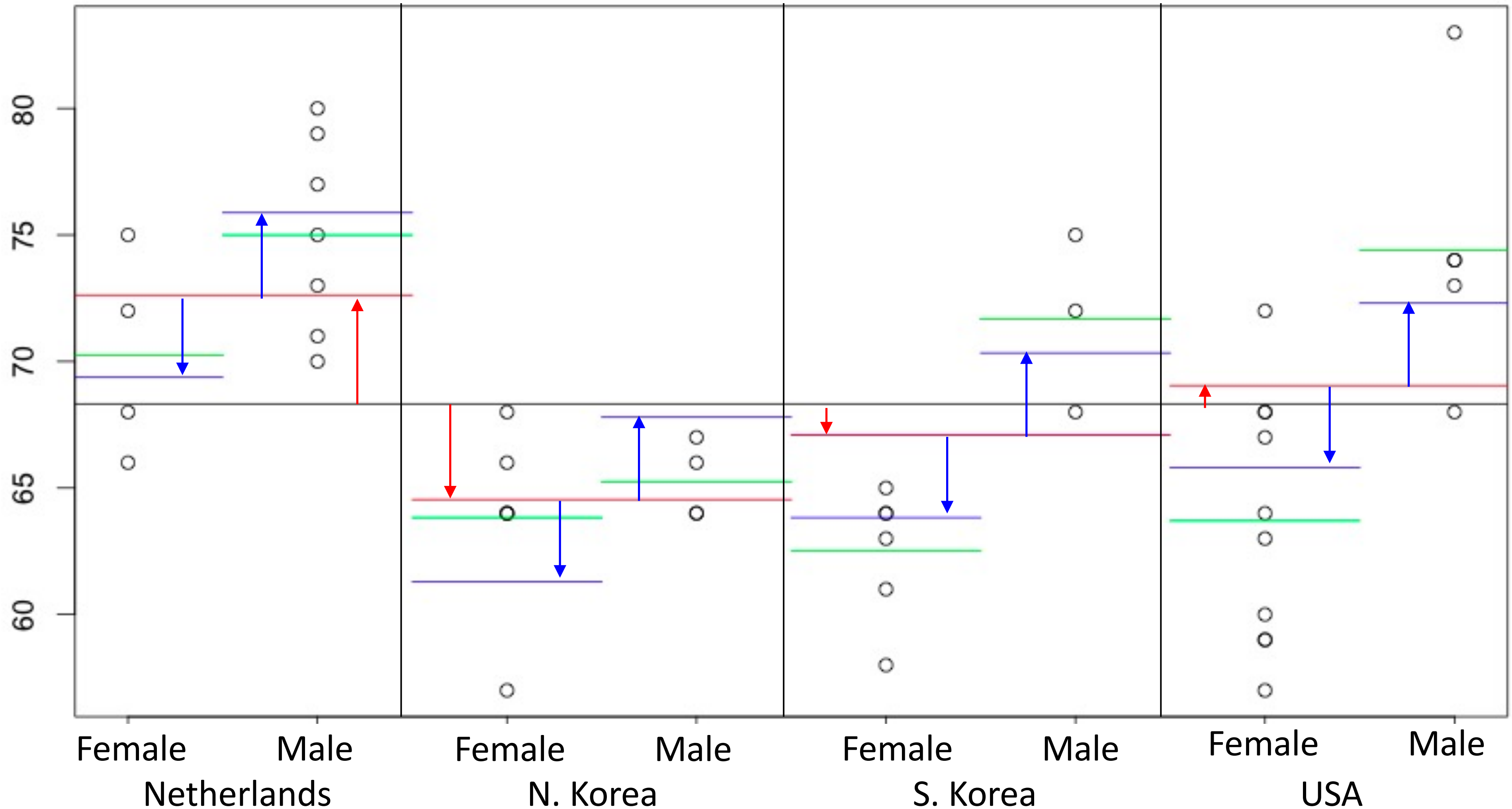
```
summary(lm(height~country+sex+country:sex))
```

```
Coefficients:
                        Estimate Std. Error t value Pr(>|t|)
(Intercept)               59.000      1.758  33.570   <2e-16 ***
countryNetherlands         3.667      2.380   1.541   0.1329
countryS.Korea             2.800      2.486   1.127   0.2681
countryUSA                 6.000      2.380   2.521   0.0167 *
sexm                       4.250      2.636   1.612   0.1165
countryNetherlands:sexm    3.917      3.478   1.126   0.2683
countryS.Korea:sexm        2.350      3.623   0.649   0.5211
countryUSA:sexm           -2.000      3.659  -0.547   0.5883
```

Interpreting coefs with interactions

This means that...

- Main effect + intercept codes for mean of cell at first level of the other factor:
  e.g., Intercept = mean of female australians
  e.g,. Intercept + B_male = mean of male australians
  e.g., Intercept + B_canada = mean of female canadians
- Interaction coefficients code for the difference unaccounted for by the 2+ levels of factors
  e.g., B_male:canada = mean(male canadians) – intercept – B_male – B_canada
- Consequently, to estimate the net effect of maleness, you have to consider both the B_male coefficient and the various B_male:country interaction terms.
  (this is something we will do more effectively with contrasts)
- Moreover, the main effect coefficients estimated without an interaction will differ from those with the interaction.



So the treatment 'main effects' are offsets for each treatment 'level' that are constant for all conditions at that treatment level and additive across factors.

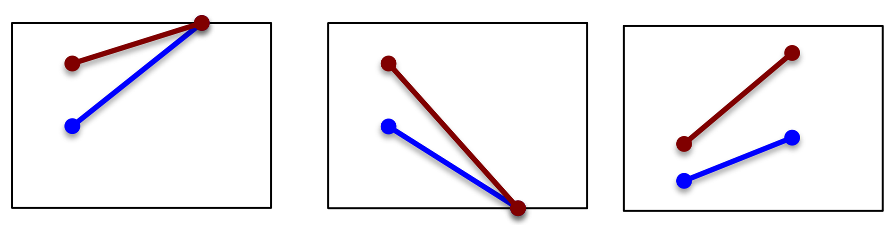
But they don't necessarily match the cell means. The distance left over is the "interaction".

What does a sig. interaction mean?

1. The variables coding for interaction account for more variance than expected by chance.
2. The additive main effects alone fail to capture variation in cell means.
3. Cell means deviate from sum of main effects.

What does it not mean?

– Effect of factor levels changes with levels of other factor. (consider ceiling, floor effects and other non-linearities)
– Means, differences, and differences of differences are what you expected.

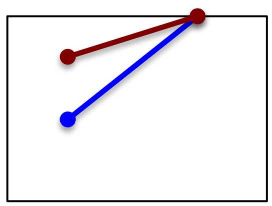


What does a sig. interaction mean?

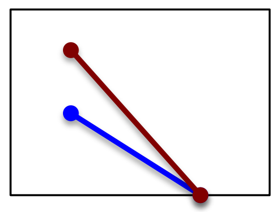
- Interaction: Main effects don't sum linearly.
- Why?
  - Influence of factor A on response variable differs in some interesting way over levels of factor B.
    eg: Major influences income only for the not rich.

What does a sig. interaction mean?

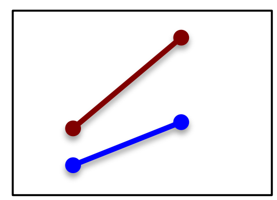
- Interaction: Main effects don’t sum linearly.
- Why?
  – Influence of factor A on response variable differs in some interesting way over levels of factor B.
  – Response variable or factor effects are not linear…
    - Ceiling effects



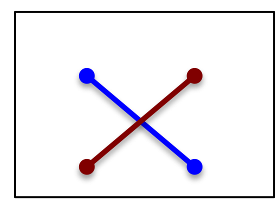
    - Floor effects



    - Multiplicative effects
    - Etc.



  – For this reason, crossover interactions are the gold standard: they rule out many non-linearities.



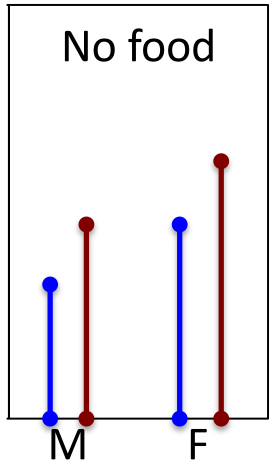
Interactions

- So what's an 'interaction'?
  - There is a difference of differences.
    e.g., the difference between male and female heights varies across countries.
  - The effect of one factor is different for different levels of an orthogonal factor.
  - More generally: influence of predictive variables (factors) on the measured variable is not additive.

Interactions

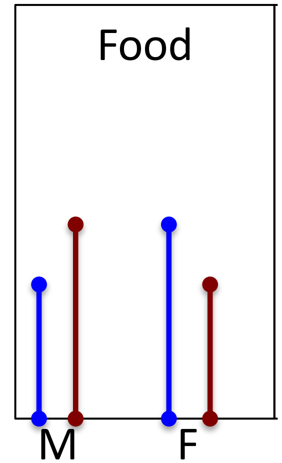
Sleepy
Awake

Two main effects,
No 2-way
interaction

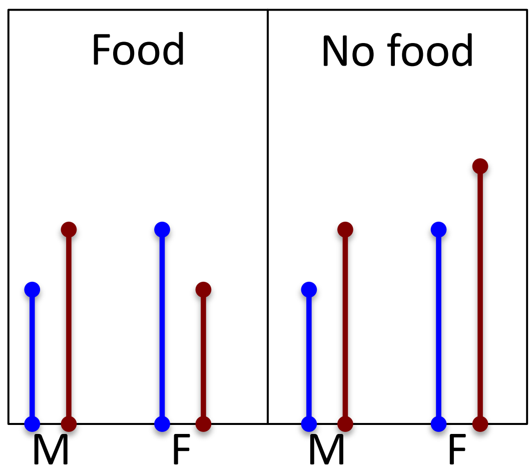


Sleepy
Awake

No main effects,
2-way ‘cross over’
interaction



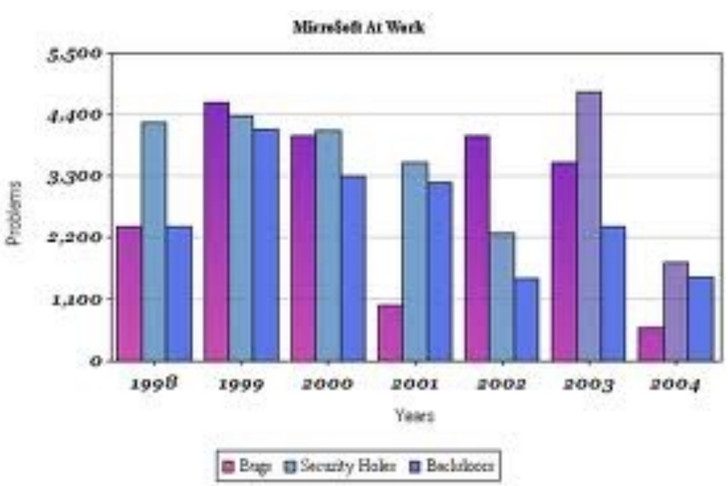
Sleepy
Awake

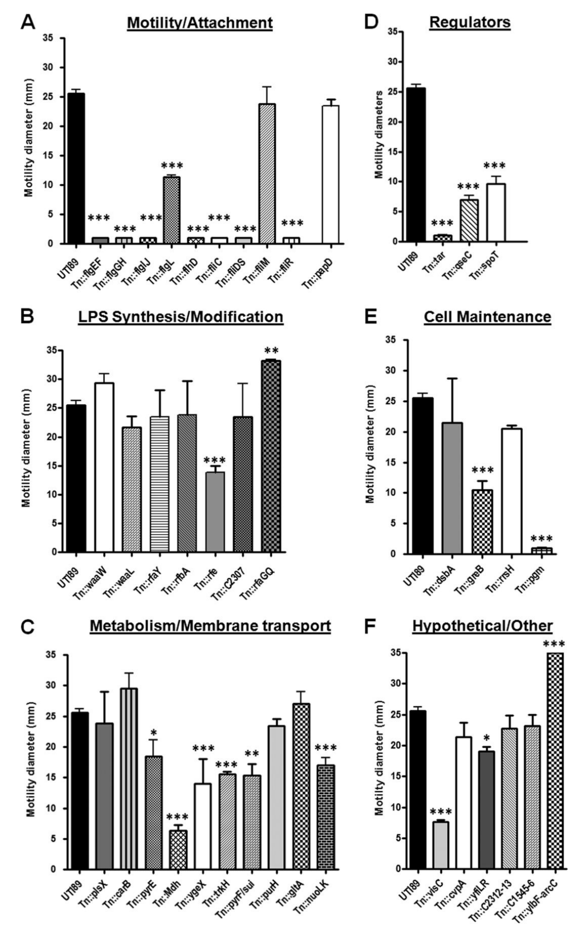


3-way interaction

Showing an interaction

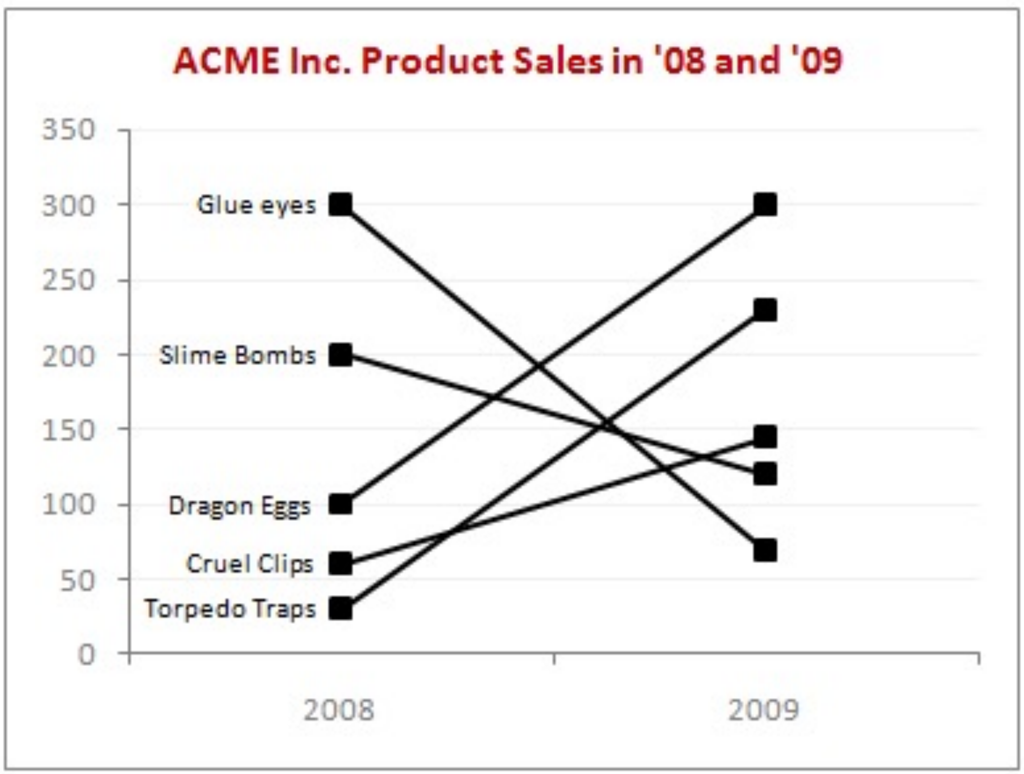
- Option 1: Bar graphs
  - Factor A: Different bars.
  - Factor B: Different groups of bars
  - Factor C: yet another grouping, or a new plot.
  - Factor D: ???
  - Factors often collapsed for display.

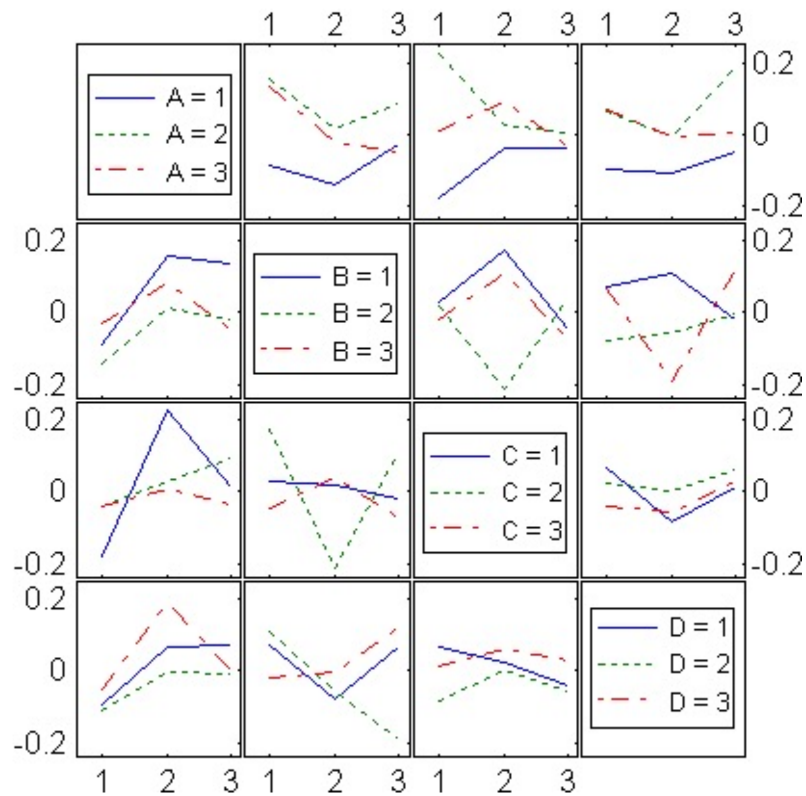




Showing an interaction

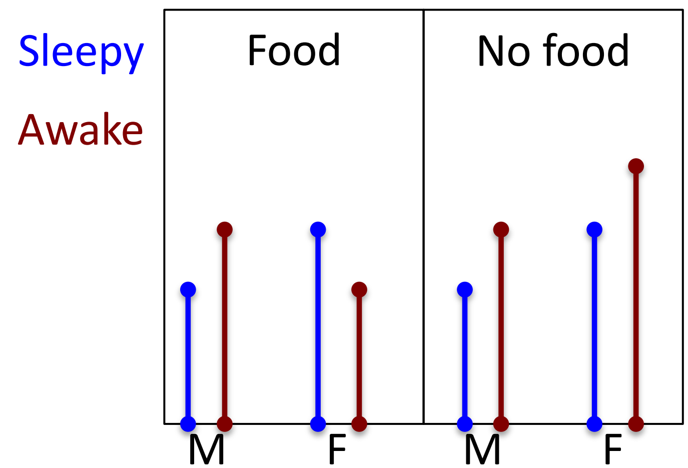
- Option 2: Line graphs
  - Factor A: different points on x axis.
  - Factor B: different lines.
  - Factor C: different panels
  - Factor D: another dimension for different panels



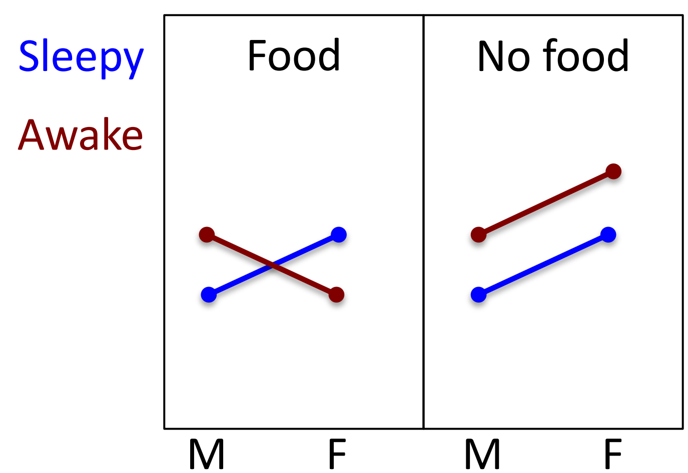


Showing an interaction

- Option 1: Bar graphs
  - + Very common!
  - + Easy to read means
  - – Wasted ink
  - – Lower data density.

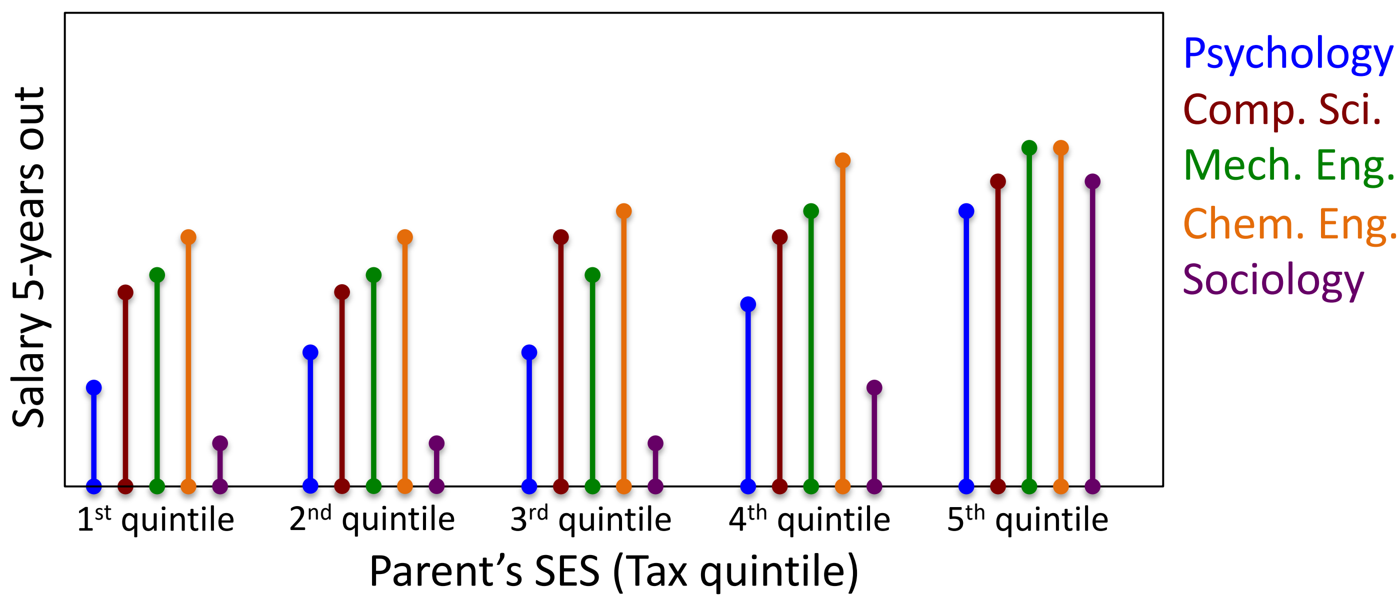


- Option 2: Line graphs
  - + High data density
  - + Easy to read interactions
  - + Less wasted ink
  - – Less common in psych.
  - + Called ‘interaction plots’ for a reason.



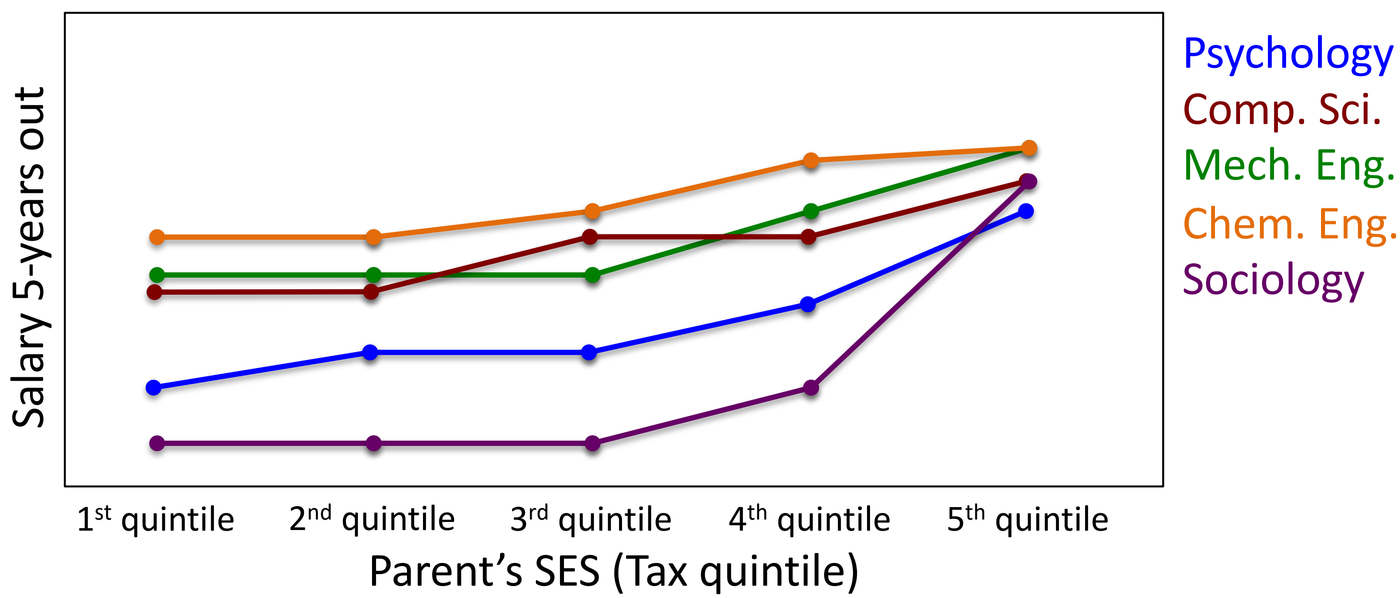
What's in these data?

- Main effect of Major?
- Main effect of Parent's SES?
- Interaction between SES and Major?



What's in these data?

- Main effect of Major?
- Main effect of Parent's SES?
- Interaction between SES and Major?



Differences of differences

- Main effect: there are differences between means of factor levels.
- 2-way interaction: the differences between means of factor A levels differ across factor B levels.
- 3-way interaction: the (differences of (differences of means of factor A levels) across factor B levels) differ across factor C levels.
- …

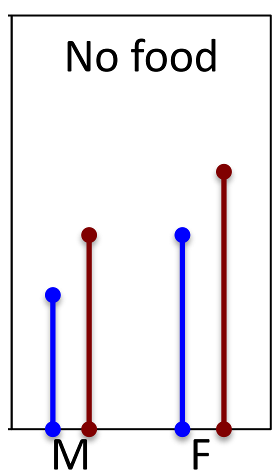
Interaction: differences

- Main effects (0^{th} order interaction?)
  - Different levels of main effect factor have different means.

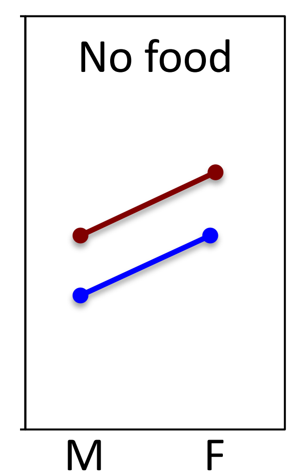
    Mean(Sleepy) < Mean(Awake)
    Mean(Male) < Mean(Female)

  - There is a difference between levels of a factor.

Sleepy
Awake

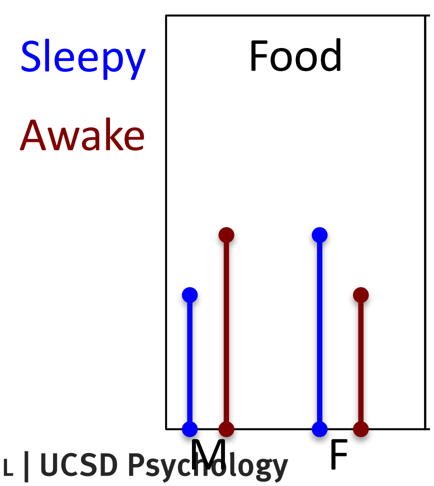


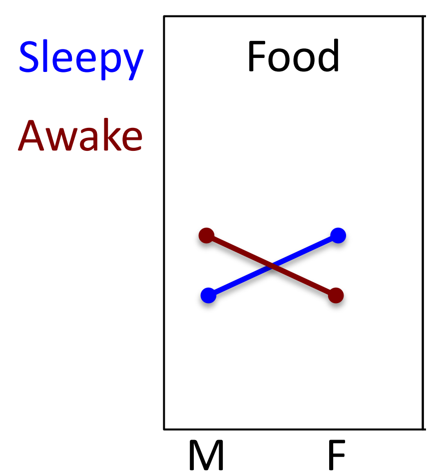
Sleepy
Awake



Interaction: differences

- 2-way Interaction (1st order interaction)
  - Differences between levels of a factor vary as a function of another factor level.
    [Mean(Sleepy|Male) – Mean(Awake|Male)]
    <
    [Mean(Sleepy|Female) – Mean(Awake|Female)]
  - There is a difference of differences.

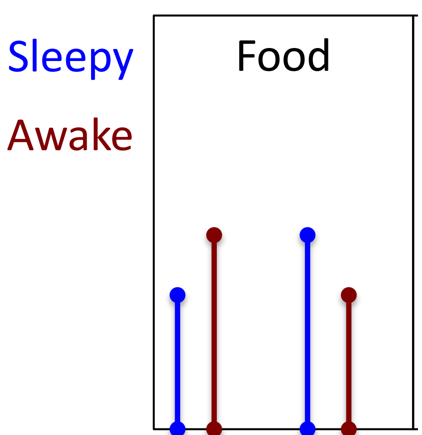


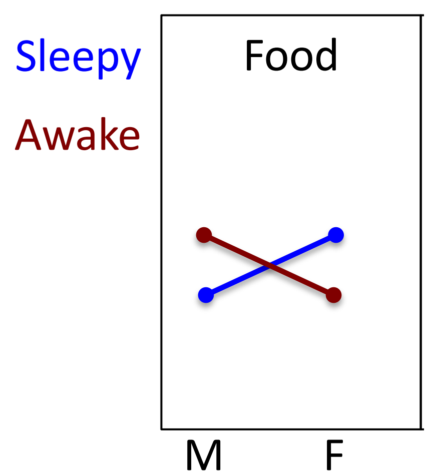


Interaction: differences

- 2-way Interaction (1st order interaction)
  - Differences between levels of a factor vary as a function of another factor level.
    [Mean(Male, Sleepy) – Mean(Female, Sleepy)]
    >
    [Mean(Male, Awake) – Mean(Female, Awake)]

  - There is a difference of differences.



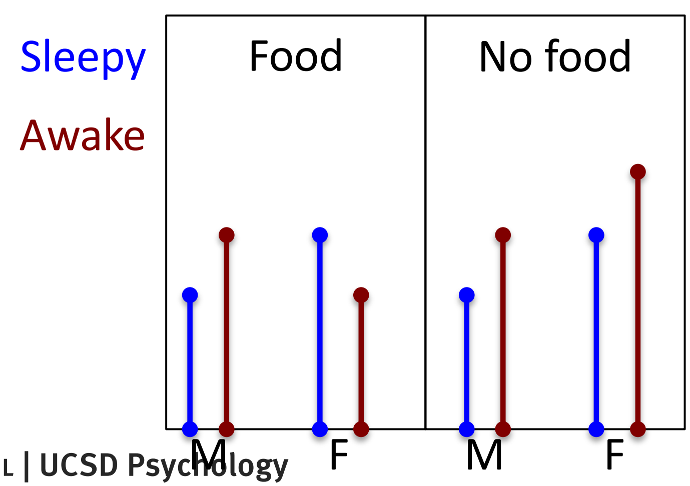


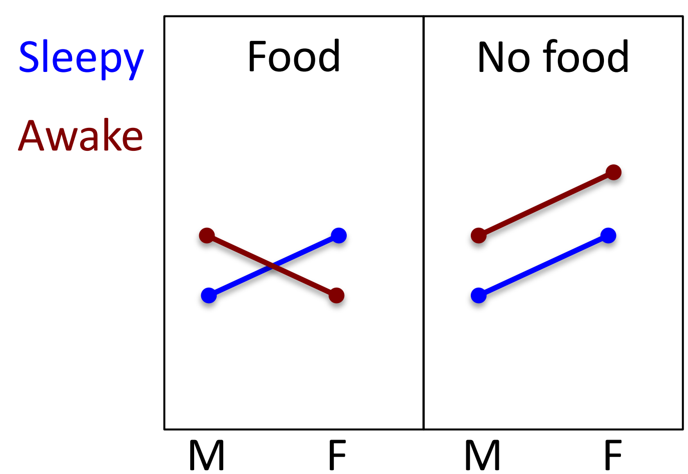
Interaction: differences

- 3-way Interaction (2nd order interaction)
  - Differences between interaction between two factors varies as a function of third-factor level.

    {[Mean(Male|Sleepy,Food) – Mean(Female|Sleepy,Food)]
    – [Mean(Male|Awake,Food) – Mean(Female|Awake,Food)]}
    >
    {[Mean(Male|Sleepy,NoFood) – Mean(Female|Sleepy, NoFood)]
    – [Mean(Male|Awake,NoFood) – Mean(Female|Awake,NoFood)]}

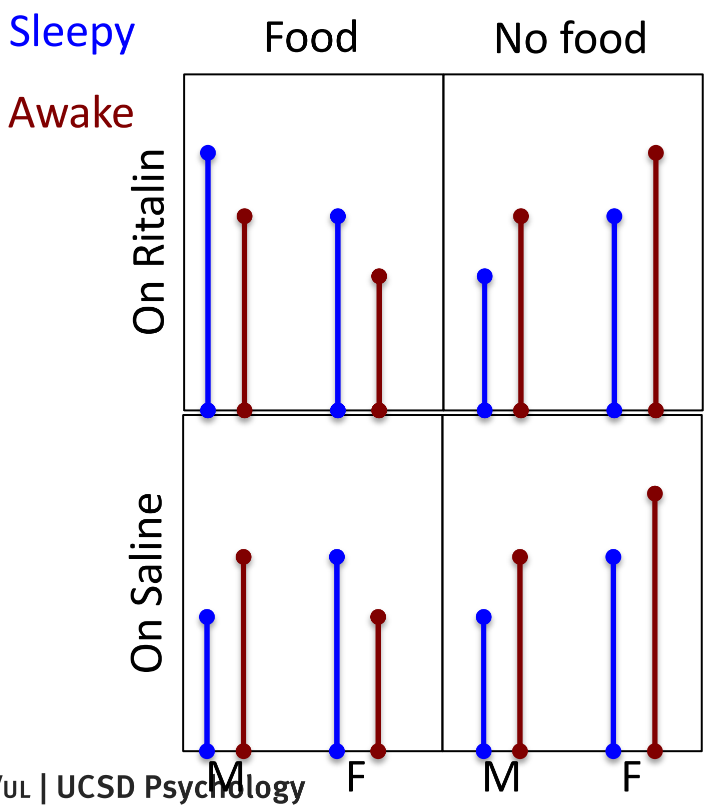
  - There is a difference of differences of differences.

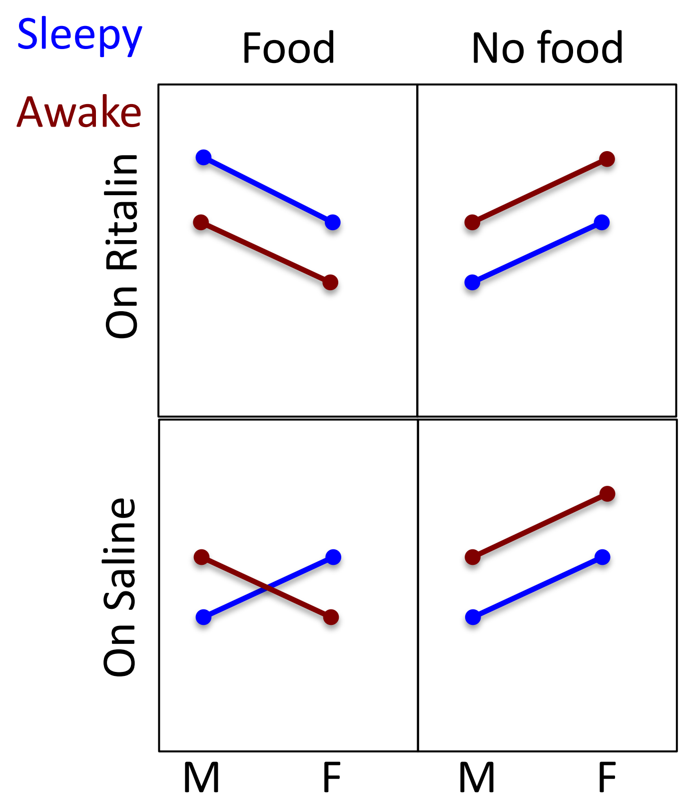




Interaction: differences

- 4-way Interaction (3rd order interaction)
  - Differences between interaction between three factors varies as a function of fourth-factor level.
  - There is a difference of differences of differences of differences.



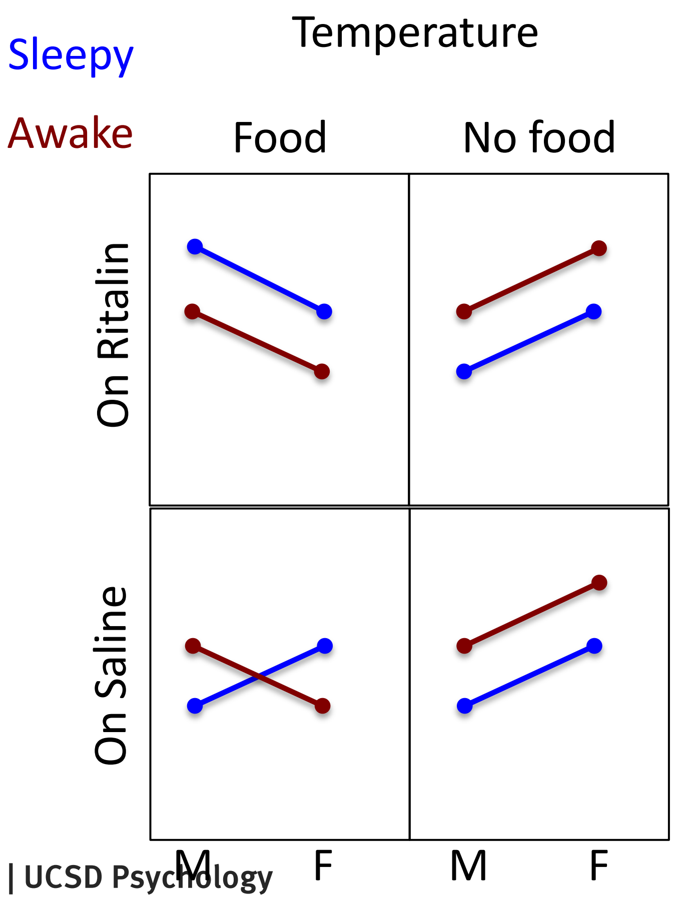


Interaction: differences

- 5-way Interaction (4th order interaction)
  - There is a difference of differences of differences of differences of differences…
  - …You get the idea… Stay away.

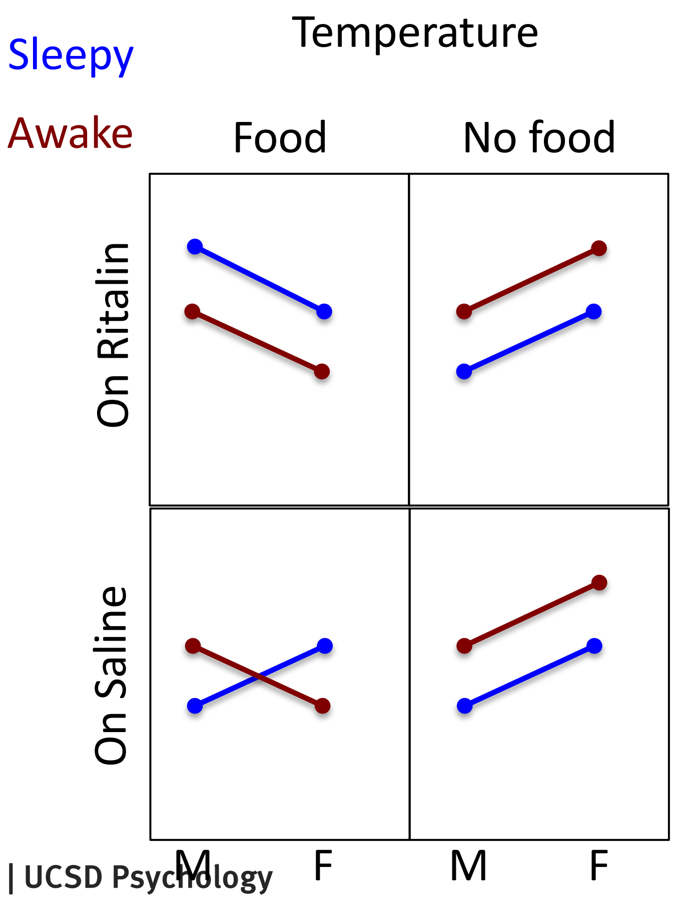
Interpreting higher order interactions via differences

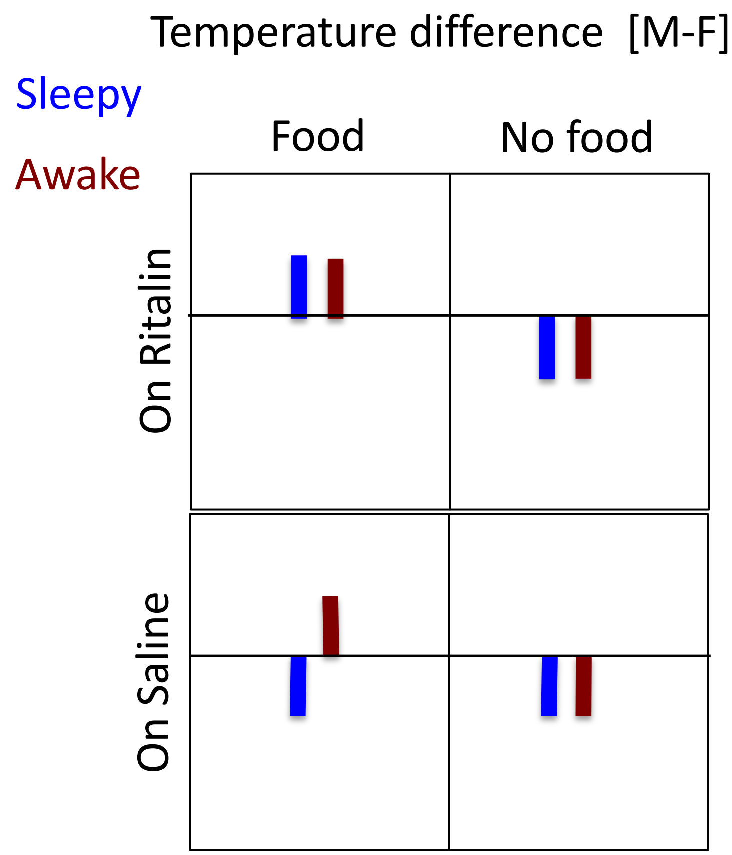
- Take the difference along one factor…



Interpreting higher order interactions via differences

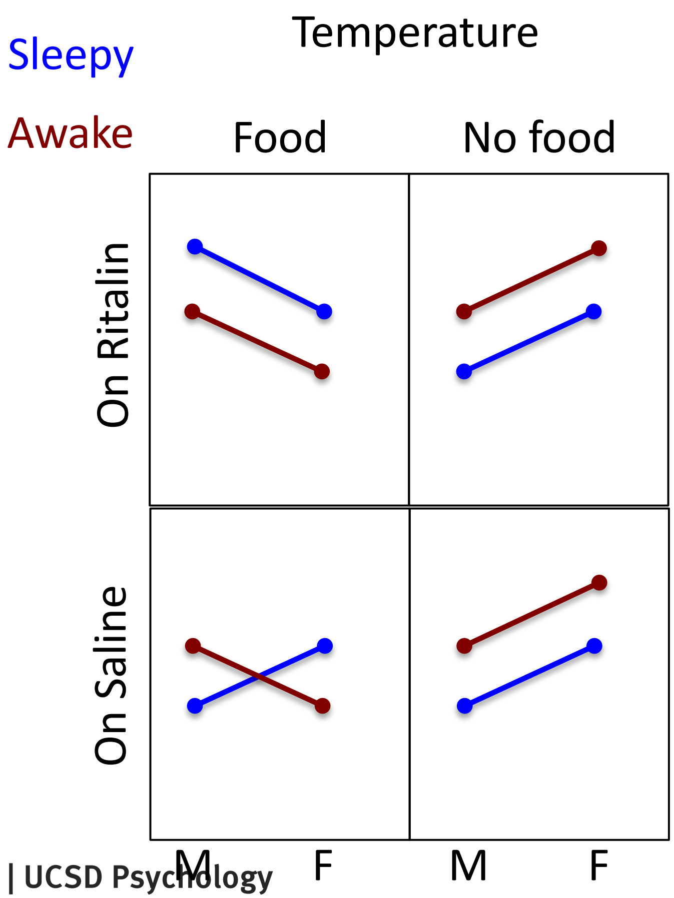
- Take the difference along one factor…

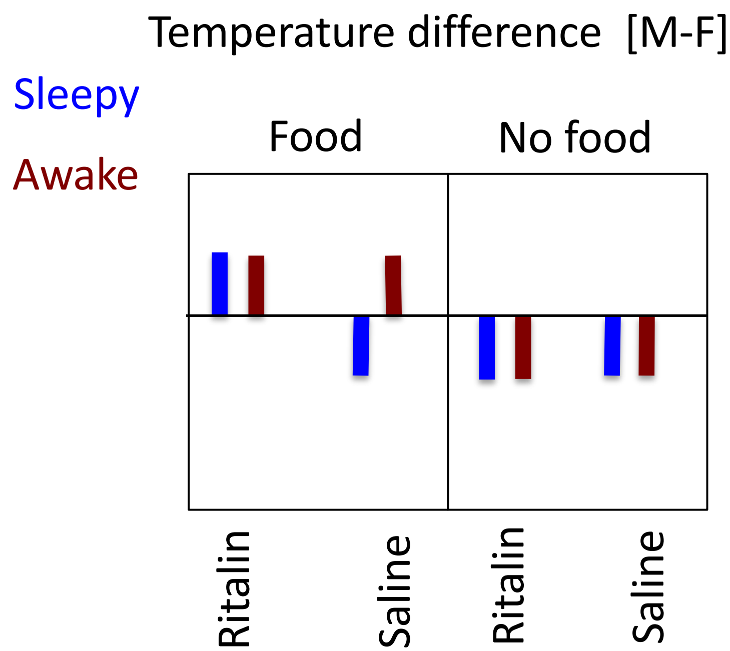




Interpreting higher order interactions via differences

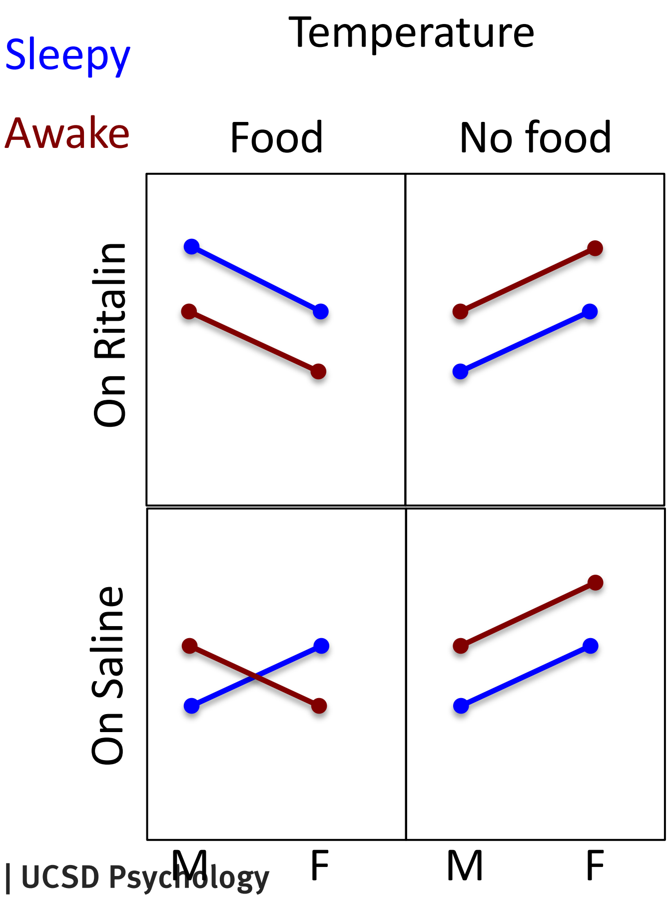
- Take the difference along one factor…

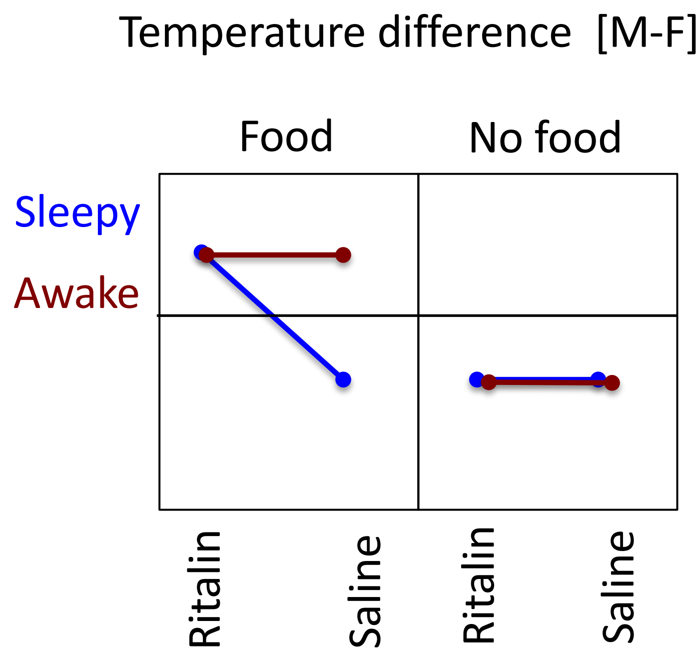




Interpreting higher order interactions via differences

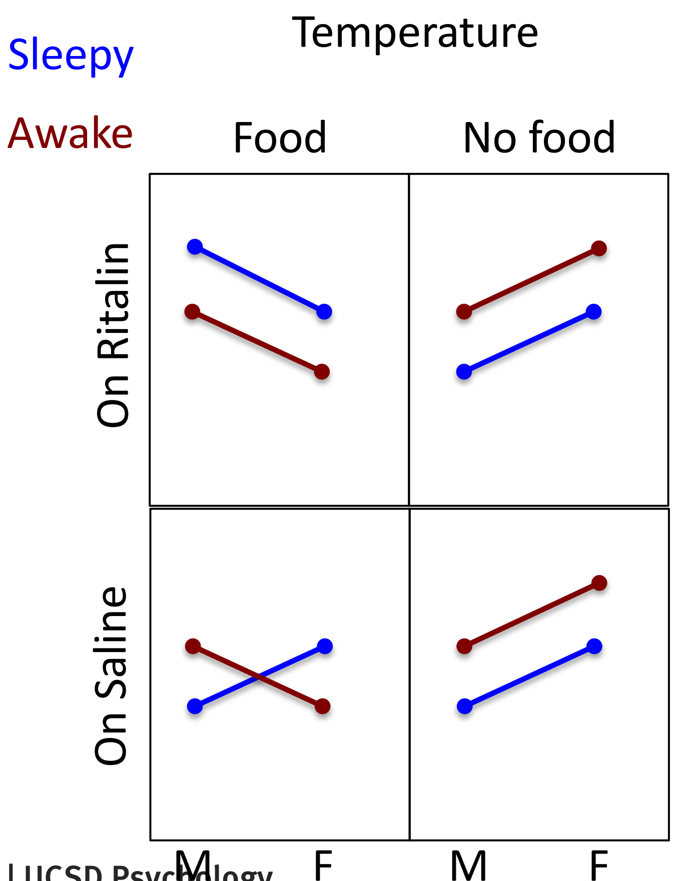
- Take the difference along one factor…

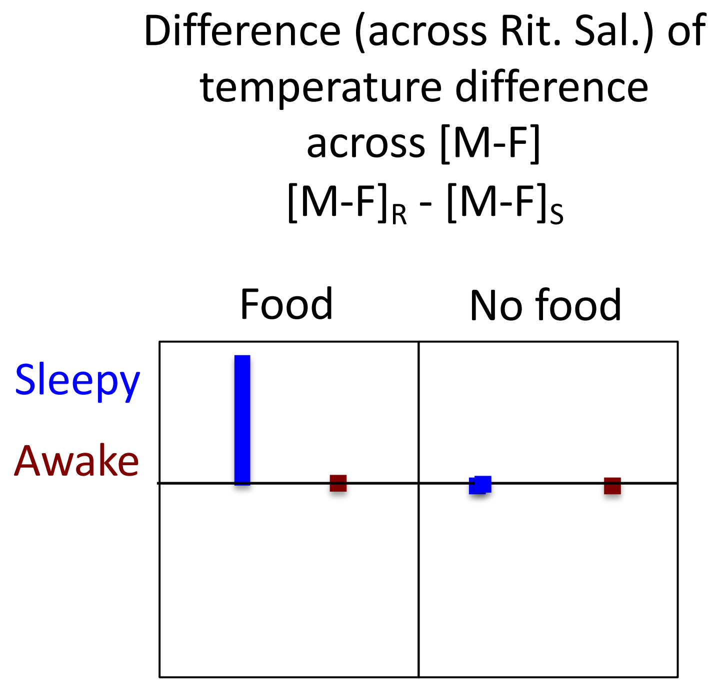




Interpreting higher order interactions via differences

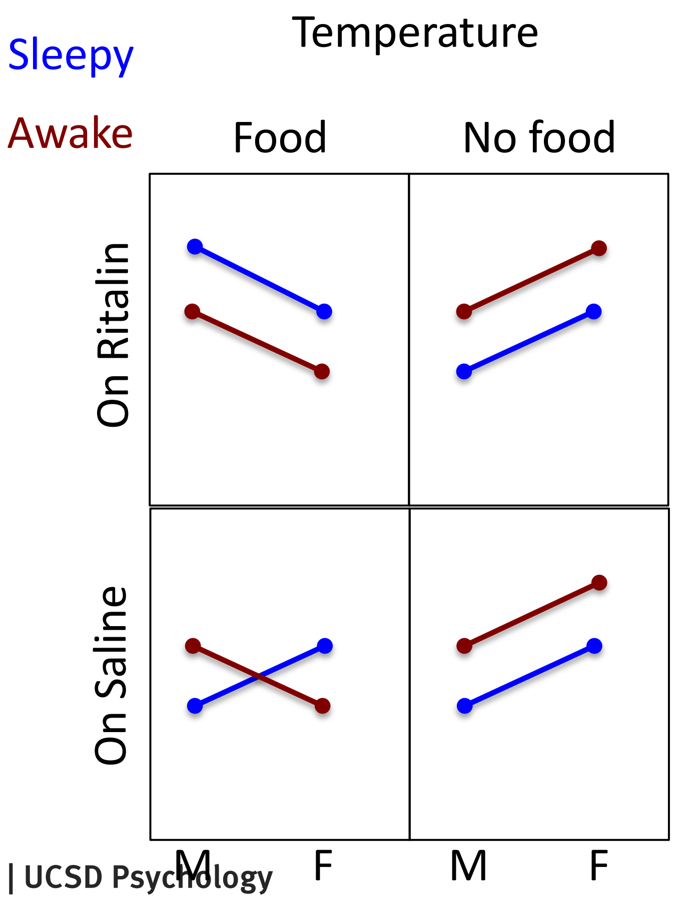
- Take the difference along one factor…





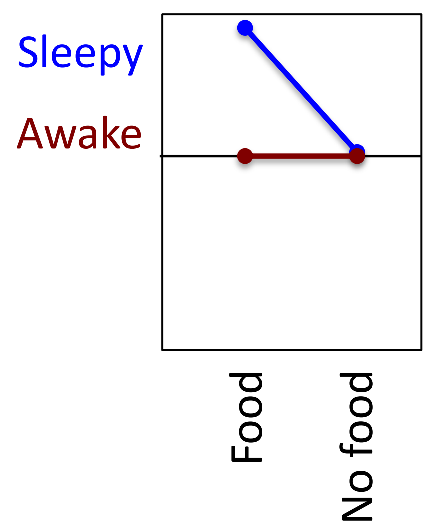
Interpreting higher order interactions via differences

- Take the difference along one factor…



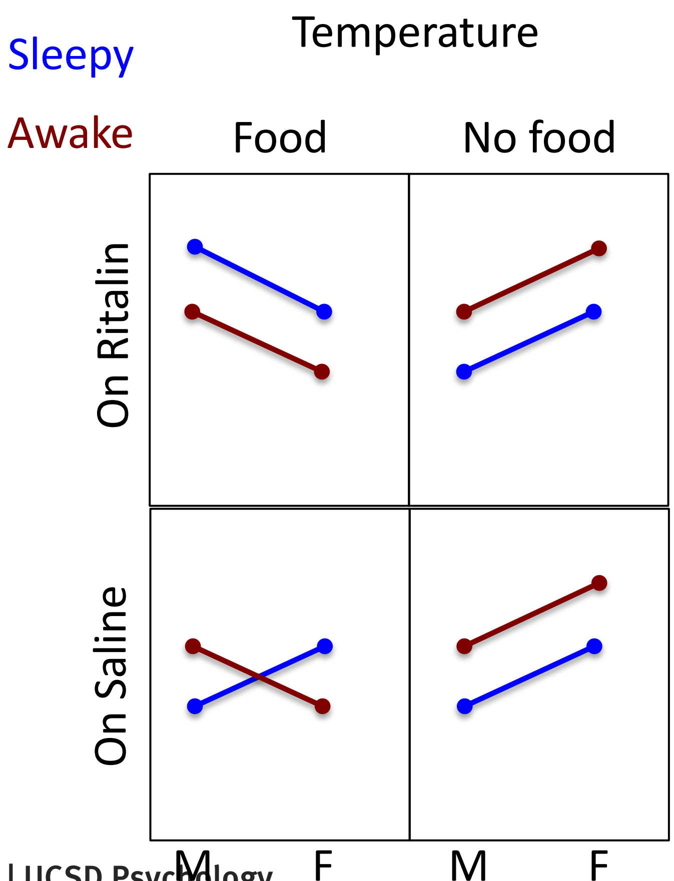
Difference (across Rit. Sal.) of temperature difference across [M-F]

$[M\text{-}F]_R - [M\text{-}F]_S$



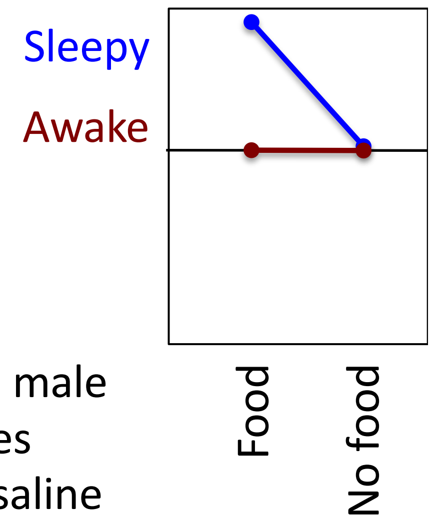
Interpreting higher order interactions via differences

- Take the difference along one factor…



Difference (across Rit. Sal.) of temperature difference across [M-F]

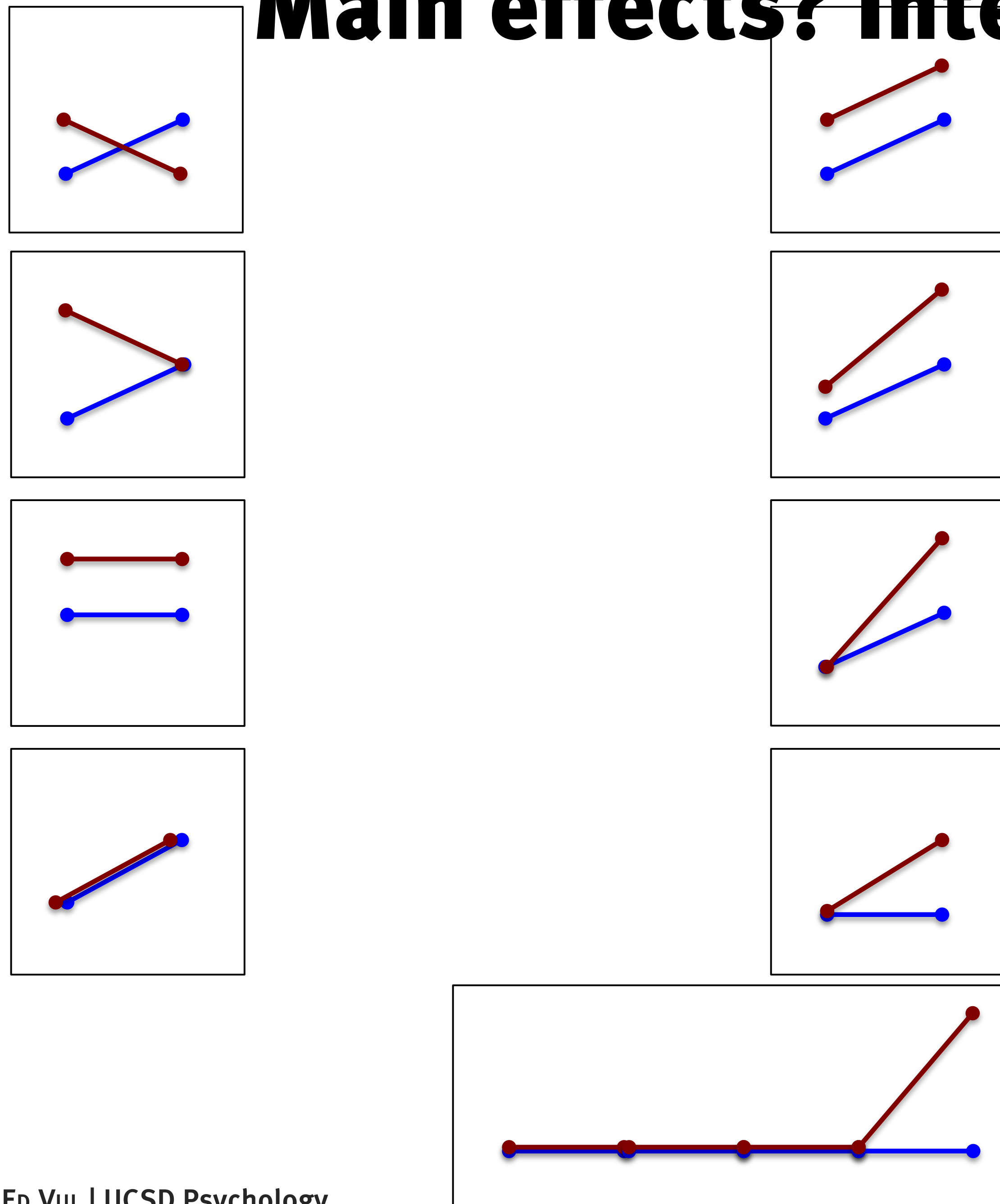
$[M\text{-}F]_R - [M\text{-}F]_S$

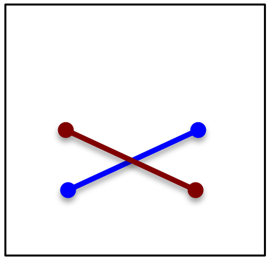


The difference between male and female temperatures differs across ritalin vs. saline but only when the hamsters are fed *and* sleepy.

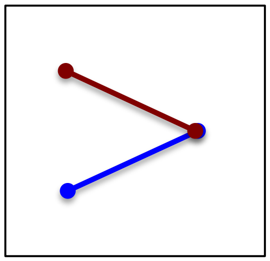
You see why higher order interactions are unwieldy…

Main effects? Interactions?

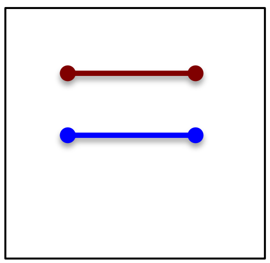




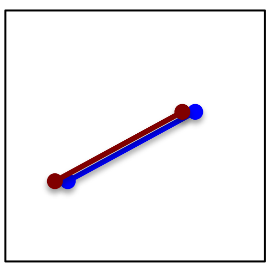
'Crossover' interaction:
No main effect of R/B
No main effect of L/R
Interaction



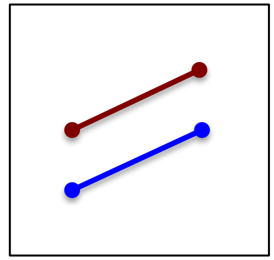
Main effect of R/B
No main effect of L/R
Interaction



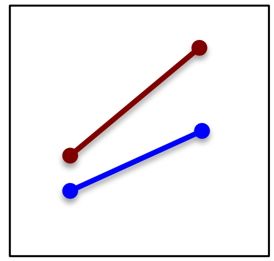
Main effect of R/B
No main effect of L/R
No Interaction



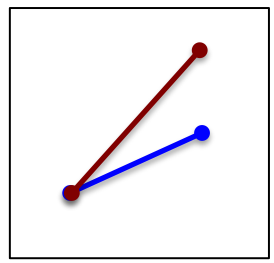
No Main effect of R/B
Main effect of L/R
No Interaction



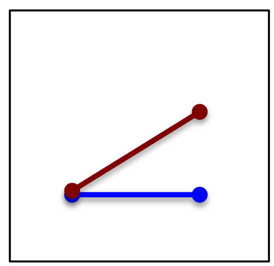
Main effect of R/B
Main effect of L/R
No Interaction



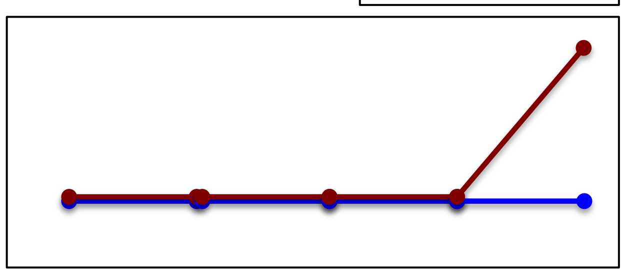
*
Main effect of R/B
Main effect of L/R
Interaction



*
Main effect of R/B
Main effect of L/R
Interaction



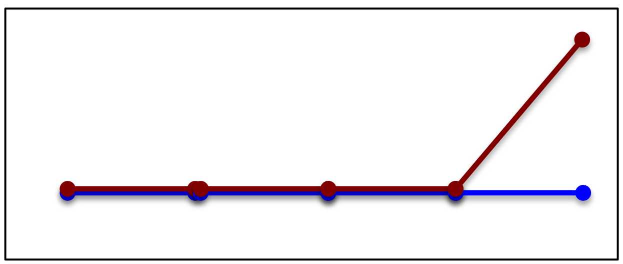
*
Main effect of R/B
Main effect of L/R
Interaction

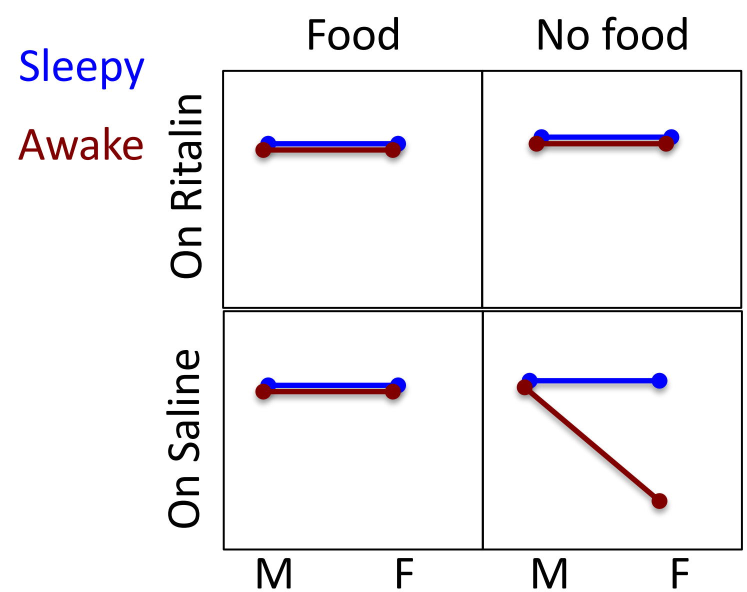


*
Main effect of R/B
Main effect of L/R
Interaction

Interactions Cautions

- Higher order interactions are hard to interpret: many (qualitatively different) patterns of means can yield the same difference of differences of differences of ….
- Main effects in the presence of an interaction (or lower order interactions in the presence of a higher order interactions) should be subject to scrutiny.

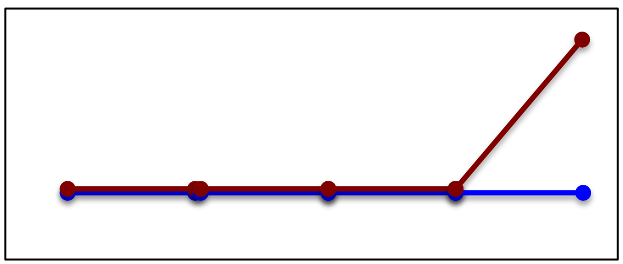


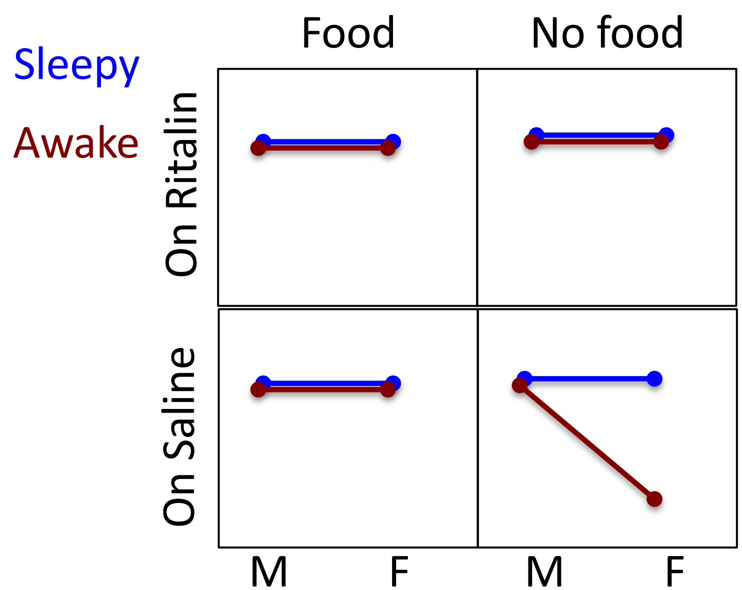


- Better to stay away from highly factorial designs unless they are strictly necessary.

Interactions Cautions

- Higher order interactions are hard to interpret: many (qualitatively different) patterns of means can yield the same difference of differences of differences of ….
- Main effects in the presence of an interaction (or lower order interactions in the presence of a higher order interactions) should be subject to scrutiny.





- Better to stay away from highly factorial designs unless they are strictly necessary.

Sums of squares in full factorial ANOVA

- SS[main effects] = sum of the squared deviations of factor level means from overall mean.
- SS[interactions] = sum of squared deviations of cell means from mean predicted by main effects.
- SS[error] = sum of squared deviations of data points from their respective cell means (deviation from predicted mean using main effects and interactions).

ANOVA table shows variance partition

```
anova(lm(height~country+sex+country:sex))
```

```
Response: height
            Df Sum Sq Mean Sq F value    Pr(>F)
country      3 196.18  65.394  4.2342   0.01226 *
sex          1 308.09 308.095 19.9486 8.803e-05 ***
country:sex  3  53.18  17.726  1.1477   0.34436
Residuals   33 509.67  15.444
```

Type I (sequential) Sums of squares: (default in R)

How much variance can country explain?	**SSR(country)**
How much more variance can sex explain?	**SSR(sex | country)**
How much more variance can the interaction explain?	**SSR(sex:country | sex, country)**

Consequently, order of factors will matter if the design is not perfectly balanced.

Type II SS: SSR(country | sex), SSR(sex | country), SSR(sex:country | sex, country)

Type III SS: SSR(country | sex, sex:country), SSR(sex | country, sex:country), SSR(sex:country | sex, country)

Type I, II, III sums of squares make different comparisons, and thus are testing different null hypotheses. Which is more appropriate depends on your question.

Degrees of freedom

- How many regressors does it take to capture a main effect?
  - # of levels minus 1
- How many regressors does it take to capture an interaction?
  - (# of levels of A minus 1)*(# of levels of B minus 1)

- think of it this way: if we code for the full model with interactions, # of parameters = # of cells (to be able to capture a unique mean for each cell).
  These get divided among intercept, main effects and interactions.

d.f. of source: Number of parameters to capture source

Sums of squares attributed to source (e.g., main effect, interaction, etc.)

$$F(\underbrace{p_{SOURCE}}_{\text{d.f. of numerator}}, \underbrace{n - p_{FULL}}_{\text{d.f. of denominator}}) = \frac{\left(\dfrac{\boxed{SSR_{SOURCE}}}{\boxed{p_{SOURCE}}}\right)}{\left(\dfrac{\boxed{SSE_{FULL}}}{\boxed{n - p_{FULL}}}\right)}$$

Residual sum of squares in full model

d.f. error of full model
(n - # all parameters)

Assumptions (and when stuff breaks)

Same as regression:

- Errors are independent…
  - Violated under sequential / temporal dependence, non-random sampling, etc.
    - Consider: mixed effects, covariates
- …identically distributed…
  - Violated if some conditions have higher variance.
    - Consider: ignoring (if not that different)
    - Consider: log transform (if errors are multiplicative)
- …and Normal.
  - Violated if measure has high skew, kurtosis, floor, ceiling effects.
    - Consider: various transformations.

Multicolinearity in unbalanced designs

	North Korea	USA
Male	67 66 64 64 68 67 69 70 65	74 83
Female	64 68	59 63 68 60 64 67 62 59 68 69

Unbalanced design: different ns in different cells, so factors are not independent, so we have multicolinearity, and a credit assignment problem.

Multicolinearity effects: Contamination across main effects, and order-dependence in sum sq. allocation.

Type I sums of squares (R default)

SS for factor 1: SSR[factor1]

SS for factor 2: SSR[factor2 | factor 1]

```
anova(lm(height~country+sex))

Response: height
          Df Sum Sq Mean Sq F value    Pr(>F)
country    3 196.18  65.394  4.1827   0.01223 *
sex        1 308.09 308.095 19.7060 8.217e-05 ***
Residuals 36 562.84  15.635
```

SSR[country] and SSR[sex|country]

```
anova(lm(height~sex+country))

Response: height
          Df Sum Sq Mean Sq F value  Pr(>F)
sex        1 316.23  316.23 20.2265 6.9e-05 ***
country    3 188.05   62.68  4.0092 0.01465 *
Residuals 36 562.84   15.63
```

SSR[sex] and SSR[country|sex]

Type II and III sums of squares, calculate SS for a given factor controlling for other stuff. II and III do not depend on order, but also don't preserve the SST = sum(all SS). Type III is default in SPSS. They implicitly test slightly different null hypotheses.

Need for contrasts…

- For designs of any sort of complexity, we often are interested in *specific patterns* of differences, not just the presence of *some* differences.

- To test for these specific patterns, we need contrasts. We will deal with those in 201b.

One observation per cell.

	North Korea	USA
Male	67	74
Female	64	59

- If we have one observation per cell, the interaction *is* the error.
- Therefore, if we include interaction in the model, we have no error left over (data points do not deviate at all from cell means).
  - Also n = # of parameters… so df error is 0…
- So we can't compute any F ratios or ascertain significance.
- Solution: omit interaction term, then that variance will be error, and you can assess main effects.

ANOVA effect size

Percent variance accounted for....

- Counterpart of R^2: η^2 "eta squared"

$$\eta_A^2 = \frac{SS[A]}{SST}$$

$$\eta_A^2 = \frac{494.57}{1716.3} = 0.288$$

Source	df	SS	MS	F	p
A (Country)	3	494.57	164.86	10	<0.001
B (Gender)	1	469.80	469.80	28.5	<0.001
A*B (Country*Gender)	3	142.14	47.38	2.87	0.049
Residuals	37	609.8	21.98		
Total	44	1716.3	25.69		

Note that this is equal to full-model R^2 when there is only one factor, but if there is more than one, it will be smaller.

ANOVA effect size

Percent variance accounted for….

- Counterpart of R^2:
  η^2 "eta squared"

$$\eta_A^2 = \frac{SS[A]}{SST}$$

$$\eta_A^2 = \frac{494.57}{1716.3} = 0.288$$

Source	df	SS	MS	F	p
A (Country)	3	494.57	164.86	10	<0.001
B (Gender)	1	469.80	469.80	28.5	<0.001
A*B (Country*Gender)	3	142.14	47.38	2.87	0.049
Residuals	37	609.8	21.98		
Total	44	1716.3	25.69		

- Partial η^2 (this is like "R^2 everything else constant")

$$partial : \eta_A^2 = \frac{SS[A]}{SS[A] + SS[error]}$$

$$partial : \eta_A^2 = \frac{494.57}{494.57 + 609.8} = 0.448$$

ANOVA effect size

Percent variance accounted for….

- Counterpart of R^2: proportion of all variance
  η^2 "eta squared"

$$\eta_A^2 = \frac{SS[A]}{SST}$$

- Counterpart of partial R^2 : "R^2 everything else constant"
  Partial η^2

$$partial: \eta_A^2 = \frac{SS[A]}{SS[A] + SS[error]}$$

But these measures are not good estimates of the effect size in the population – they are biased because SS[A] includes some variance due to noise…

ANOVA effect size.

- There is a surprisingly large number of candidate effect sizes for an ANOVA, all interrelated, but with slightly different properties.
  - η^2, ω^2, f^2, f, Ψ, …
- What do we want from an effect size?
  - Quantify standardized relationship strength in population (independence from sample size)
  - …in an interpretable way
  - …that we can estimate from a sample
  - …and will allow us to predict power
  - …while generalizing across study designs

My preference: ω^2 (omega squared)

- Effect size: Variance of signal in population, relative to unexplained variance in population.

$$\omega^2_{Source} = \frac{\sigma^2_{Source}}{\sigma^2_{Source} + \sigma^2_{Error}}$$

- It's like partial η^2, but is a population property
  - So to generalize across designs, it must assume that variability due to other factors was introduced by the experiment, and will not occur otherwise.
- Partial η^2 overestimates; we need a correction.

$$\hat{\omega}^2_{Source} = \frac{SS[Source] - df_{source} \cdot MS[Error]}{SS[Source] + (N - df_{source}) \cdot MS[Error]}$$

ω² and other measures

$$f^2_{Source} = \frac{\omega^2_{Source}}{1-\omega^2_{Source}} = \frac{\sigma^2_{Source}}{\sigma^2_{Error}}$$

This is a "signal-to-noise" ratio measurement: Variance of signal divided by variance of noise.

$$f_{Source} = \sqrt{\frac{\omega^2_{Source}}{1-\omega^2_{Source}}} = \frac{\sigma_{Source}}{\sigma_{Error}}$$

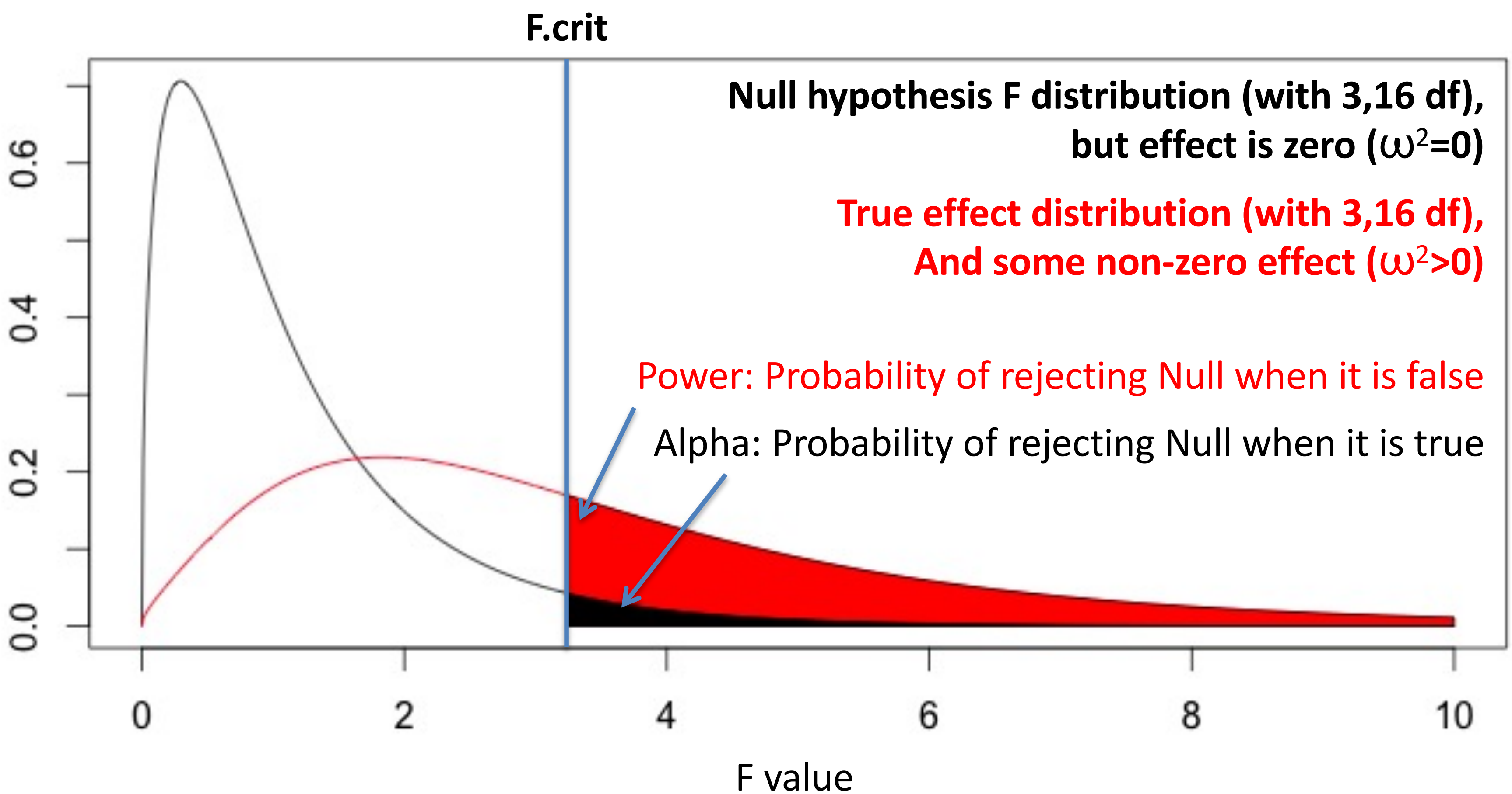
This is a "signal-to-noise" ratio measurement in original (not squared) units, thus is more analogous to Cohen's d

$$\lambda = \mathrm{N} * f^2_{Source} = N * \frac{\omega^2_{Source}}{1-\omega^2_{Source}}$$

This is the F distribution "non-centrality parameter" used to describe the distribution of F statistics obtained when samples come from a distribution with some real effect.

What's a big effect? Some say ω^2=0.15 is big, 0.06 is medium, 0.01 is small.

Power for the F-test



So, to figure out the power of an F test we need to know the sample size, alpha, and true effect.

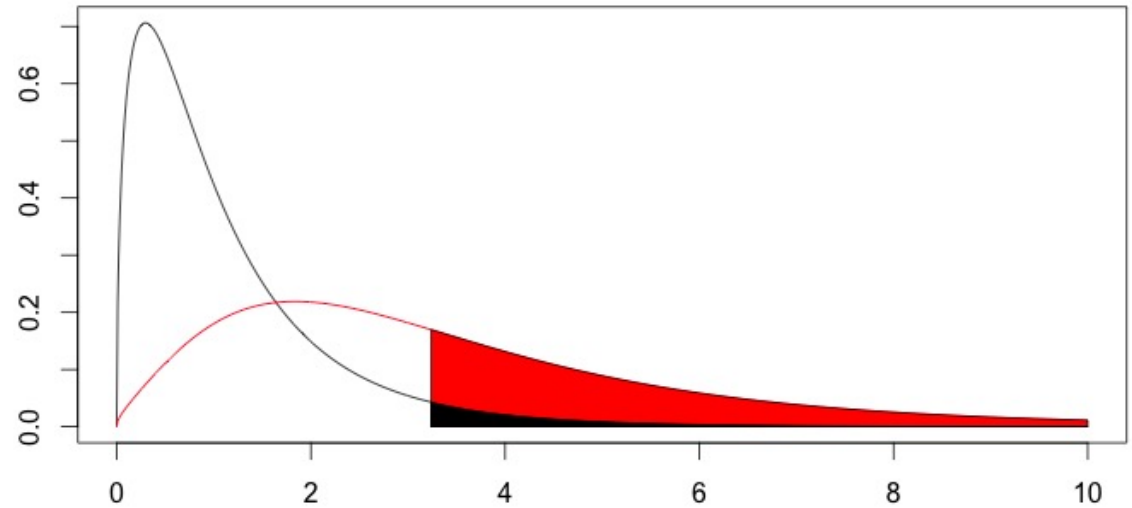
Power for the F-test

Total number of cells `k=4`

Total (balanced) sample size `N = k*10`

Effect size (ω^2) `w2 = 0.25`

alpha `alpha = 0.05`



Non-centrality parameter

```
lambda = N*w2/(1-w2)
```
```
[1] 13.33
```

F value at which we reject H0

```
f.crit = qf(1-alpha, k-1, N-k)
```
```
[1] 2.866266
```

Power

```
power = 1-pf(f.crit, k-1, N-k, lambda)
```
```
[1] 0.84
```

Required n for certain power

This is trickier, as changing n changes both the null distribution and the true-effect distribution

```
n = 5
power = 1-pf(qf(0.95, k-1, k*(n-1)), k-1, k*(n-1), n*k*w2/(1-w2))     [1] 0.46

n = 6
power = 1-pf(qf(0.95, k-1, k*(n-1)), k-1, k*(n-1), n*k*w2/(1-w2))     [1] 0.56

n = 7
power = 1-pf(qf(0.95, k-1, k*(n-1)), k-1, k*(n-1), n*k*w2/(1-w2))     [1] 0.65

n = 8
power = 1-pf(qf(0.95, k-1, k*(n-1)), k-1, k*(n-1), n*k*w2/(1-w2))     [1] 0.73

n = 9
power = 1-pf(qf(0.95, k-1, k*(n-1)), k-1, k*(n-1), n*k*w2/(1-w2))     [1] 0.79

n = 10
power = 1-pf(qf(0.95, k-1, k*(n-1)), k-1, k*(n-1), n*k*w2/(1-w2))     [1] 0.84

n = 11
power = 1-pf(qf(0.95, k-1, k*(n-1)), k-1, k*(n-1), n*k*w2/(1-w2))     [1] 0.88
```

So we have to solve for it numerically… I recommend using the pwr R package.

Drawing data consistent with ANOVA

1) The San Diego K-12 Education board is trying to evaluate the efficacy of their math teachers. They measure average pre-to-post class improvement on a standardized test for different teachers, as a function of teacher seniority (years teaching: 0-5, 5-10, 10-15, 15-20, 20+), teacher gender (male, female), and teacher college major (STEM, Humanities, Social Science). Their analysis reveals no main effect of seniority, no main effect of gender, a significant main effect of major (STEM > Social Science > Humanities), and a significant interaction between a quadratic trend for seniority and gender. No other effects were found. Draw plot(s) showing a pattern of means that would be consistent with these effects.

ANOVA table sudoku

Length of prison sentence was measured as a function of Crime (3 levels: theft, fraud, arson) and Time of day that the judge made the decision. (5 levels: 8-9:30, 9:30-11, 11-12:30, 1:30-3, 3-4:30)

3a. Fill in the blanks given that there are five observations per condition. (write p to 3 sig digits)

Source	SS	df	MS	F	p
Crime	45	_____	_____	_____	_____
Time	85	_____	_____	_____	_____
Crime*Time	120	_____	_____	_____	_____
Error	_____	_____	_____		
Total	700				

Coefficients

Length of prison sentence was measured as a function of Crime (3 levels: theft, fraud, arson) and Time of day that the judge made the decision. (5 levels: 8-9:30, 9:30-11, 11-12:30, 1:30-3, 3-4:30)

```
summary(lm(sentence.mo~crime*time))
Coefficients:
                         Estimate
(Intercept)                  60
Crime-fraud                 -12
Crime-theft                  4
Time-0930                        -3
Time-1100                         8
Time-1330                        -5
Time-1500                         6
Crime-fraud:Time-0930        0
Crime-theft:Time-0930       -3
Crime-fraud:Time-1100       +5
Crime-theft:Time-1100       -2
Crime-fraud:Time-1330       -2
Crime-theft:Time-1330        2
Crime-fraud:Time-1500       -1
Crime-theft:Time-1500        10
```

<- Made up!

What are the mean prison sentences in all 15 crime*time cells? (assuming R's default factor coding scheme)

ANOVA table sudoku

4a) You get your own data on math education teacher efficacy. You measure pre-post test improvement in 120 classes, 10 in each cell of a 3 teacher-major [STEM/humanities/social science] by 4 time-of-day [9:30am, 11am, 12:30pm, 2pm] design. Please fill in the following ANOVA table

	SS	df	MS	F	p
TeacherMajor	_______	_________	_______	**2.5**	______
TimeOfDay	**300**	_________	_______	_______	______
TeacherMajor X TimeOfDay	_______	_________	_______	_______	______
Error	_______	________	**10**		
Total	**1610**	_________			