PSYCH 201B

Statistical Intuitions for Social Scientists

Modeling data VII

You can download these slides:
course website > Week 8 > Overview

Today's Plan

- Meet the Final Project Proposal: https://tinyurl.com/201b-proposal
- Slides: Multiple Categorical Predictors
- *BREAK*
- Notebook: On your own

02/24/2025

Announcements

- Final Project Proposal

Final Project Proposal

- https://tinyurl.com/201b-proposal
- **March 12th:** proposal approval deadline
  - you must have **met with us and received approval** by this date
- **March 20th:** final project deadline; you will submit
  - Methods & Results section write-up **PDF**
  - Data Analysis **Notebook(s)**

Announcements

- Final Project Proposal
- HW 3 **Due tomorrow by 4pm**
- **No new HW** this week
- Instead, please check-out this week's readings

Readings (Monday & Tuesday's materials)

- Data Analysis: A Model Comparison Approach
  - Chapter 8: One-Way ANOVA: Models with a Single Categorical Predictor
  - Chapter 9: Factorial ANOVA: Models with Multiple Categorical Predictors and Product Terms
  - Chapter 10: ANCOVA: Models with Continuous and Categorical Predictors
- Regression and Other Stories
  - Chapter 10: Linear Regression with Multiple Predictors

Announcements

- Final Project Proposal
- HW 3 **Due tomorrow by 4pm**
- **No new HW** this week
- Instead, please check-out this week's readings
- After the BREAK we have *solution* notebooks from last week and a new notebook to complete
  - `06_models_solutions` (check your work today)
  - `07_models_solutions` (check your work today)
  - `08_models` (work through today)

Review: Categorical predictors (3+ levels)

Dataset

skill	hand	limit	balance
expert	bad	fixed	4.00
expert	bad	fixed	5.55
expert	bad	none	5.52
expert	bad	none	8.28
expert	neutral	fixed	11.74
expert	neutral	fixed	10.04
expert	neutral	none	21.55
expert	neutral	none	3.12
expert	good	fixed	10.86
expert	good	fixed	8.68

skill = expert/average

hand = bad/neutral/good

limit = fixed/none

balance = final balance in Euros

2 (skill) x 3 (hand) x 2 (limit) design

25 participants per condition

n = 300

Meyer, G., von Meduna, M., Brosowski, T., & Hayer, T. (2012). Is poker a game of skill or chance? A quasi-experimental study. *Journal of Gambling Studies*

Do better hands win more money?

skill	hand	limit	balance
expert	bad	fixed	4.00
expert	bad	fixed	5.55
expert	bad	none	5.52
expert	bad	none	8.28
expert	neutral	fixed	11.74
expert	neutral	fixed	10.04
expert	neutral	none	21.55
expert	neutral	none	3.12
expert	good	fixed	10.86
expert	good	fixed	8.68

skill = expert/average

hand = bad/neutral/good

limit = fixed/none

balance = final balance in Euros

2 (skill) x 3 (hand) x 2 (limit) design

25 participants per condition

n = 300

Meyer, G., von Meduna, M., Brosowski, T., & Hayer, T. (2012). Is poker a game of skill or chance? A quasi-experimental study. *Journal of Gambling Studies*

Do better hands win more money?

skill	hand	limit	balance
expert	bad	fixed	4.00
expert	bad	fixed	5.55
expert	bad	none	5.52
expert	bad	none	8.28
expert	neutral	fixed	11.74
expert	neutral	fixed	10.04
expert	neutral	none	21.55
expert	neutral	none	3.12
expert	good	fixed	10.86
expert	good	fixed	8.68

Regression with one **categorical predictor**

- $\text{balance}_i \sim \text{hand}_i$

We need to **encode** *k* (three) levels using *k-1* (two) parameters

- $Y_i = \beta_0 + \beta_1 \text{hand}_i + \beta_2 \text{hand}_i$

Treatment/dummy (default)

- $\beta_0 =$ reference level
- $\beta_n =$ level_n- reference level

Deviation/sum

- $\beta_0 =$ grand mean
- $\beta_n =$ level_n- grand mean

Polynomial/orthogonal

- $\beta_0 =$ grand mean
- $\beta_n =$ trend of order_{n+1}

Do better hands win more money?

skill	hand	limit	balance
expert	bad	fixed	4.00
expert	bad	fixed	5.55
expert	bad	none	5.52
expert	bad	none	8.28
expert	neutral	fixed	11.74
expert	neutral	fixed	10.04
expert	neutral	none	21.55
expert	neutral	none	3.12
expert	good	fixed	10.86
expert	good	fixed	8.68

Regression with one **categorical predictor**

- $\text{balance}_i \sim \text{hand}_i$

We need to **encode** *k* (three) levels using *k-1* (two) parameters

- $Y_i = \beta_0 + \beta_1 \text{hand}_i + \beta_2 \text{hand}_i$

```
treatment = ols('balance ~ C(hand)', data=df.to_pandas())
```

- $\beta_0 =$ reference level
- $\beta_n =$ level$_n$- reference level

```
sums = ols('balance ~ C(hand, Sum)',data=df.to_pandas())
```

- $\beta_0 =$ grand mean
- $\beta_n =$ level$_n$- grand mean

```
polys = ols('balance ~ C(hand, Poly)',data=df.to_pandas())
```

- $\beta_0 =$ grand mean
- $\beta_n =$ trend of order$_{n+1}$

One-Way ANOVA: Coding doesn't matter

```
# Compact
model_c = ols('balance ~ 1', data=df.to_pandas())
results_c = model_c.fit()

# Augmented
model_a = ols('balance ~ C(hand)', data=df.to_pandas())
results_a = model_a.fit()

# Worth it?
anova_lm(results_c, results_a)
```

Equivalent!

```
anova_lm(results_a, type=3)
```

✓ 0.0s

	df	sum_sq	mean_sq	F	PR(>F)
C(hand)	2.0	2559.401402	1279.700701	75.702581	2.699281e-27
Residual	297.0	5020.583223	16.904321	NaN	NaN

The final balance differed significantly as a function of the quality of a player's hand (i.e. whether the hand was bad, neutral, or good), **$F(2, 297) = 75.703$, $p < .001$**

So when/why does it matter?

Multiple Regression w/ Categorical Predictors (3+ levels)

Do better hands affect balance differently based on skill?

skill	hand	limit	balance
expert	bad	fixed	4.00
expert	bad	fixed	5.55
expert	bad	none	5.52
expert	bad	none	8.28
expert	neutral	fixed	11.74
expert	neutral	fixed	10.04
expert	neutral	none	21.55
expert	neutral	none	3.12
expert	good	fixed	10.86
expert	good	fixed	8.68

skill = expert/average

hand = bad/neutral/good

limit = fixed/none

balance = final balance in Euros

2 (skill) x 3 (hand) x 2 (limit) design

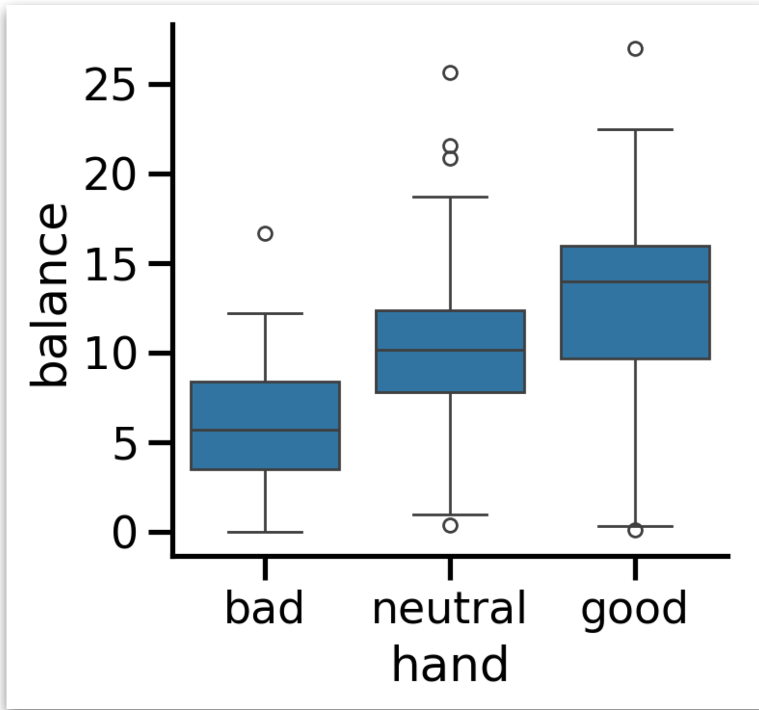
25 participants per condition

n = 300

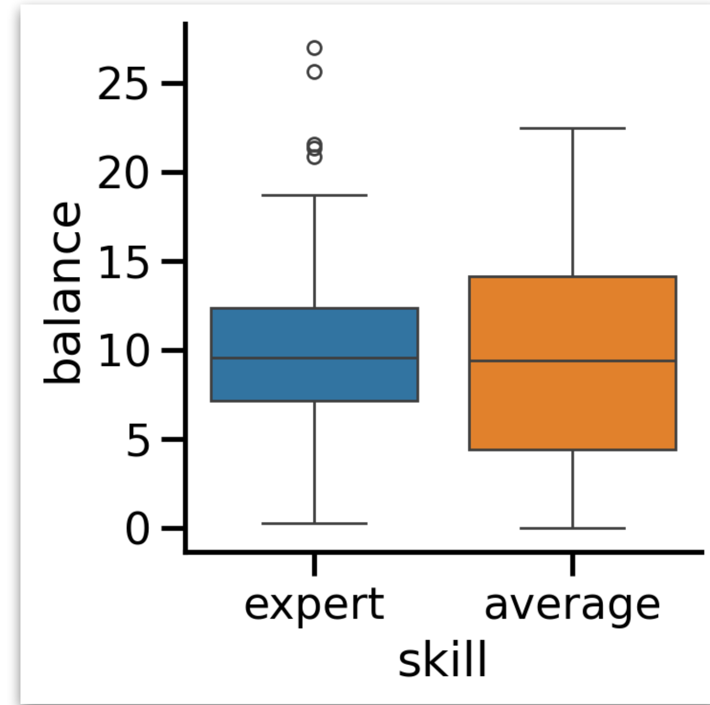
Meyer, G., von Meduna, M., Brosowski, T., & Hayer, T. (2012). Is poker a game of skill or chance? A quasi-experimental study. *Journal of Gambling Studies*

Do better hands affect balance differently based on skill?

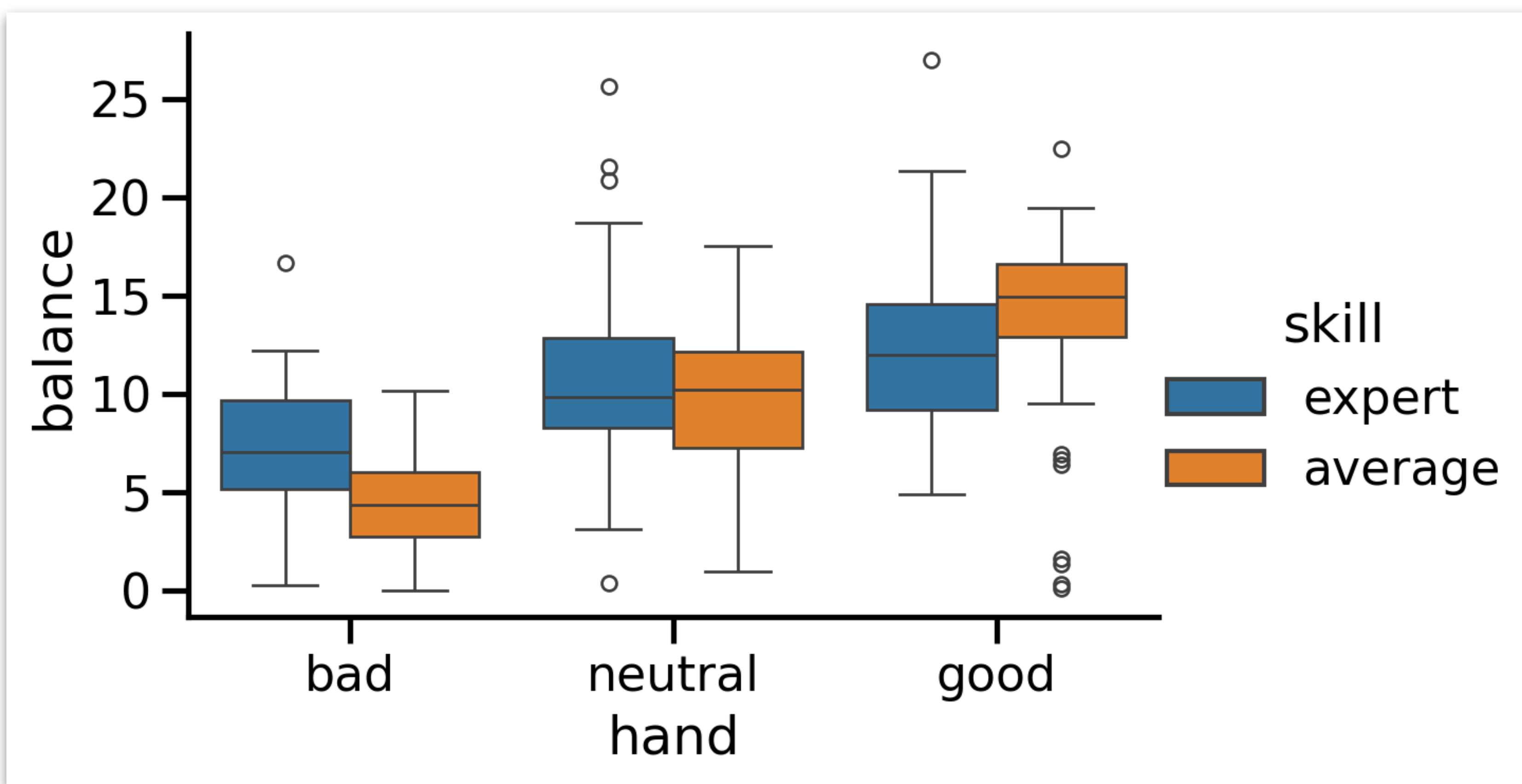
Effect of Hand



Effect of Skill

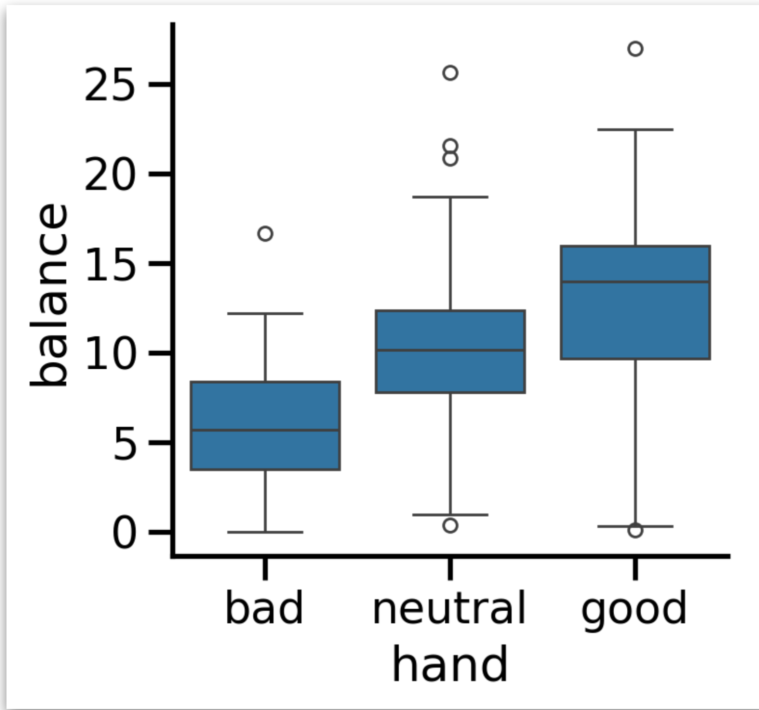


Hand & Skill

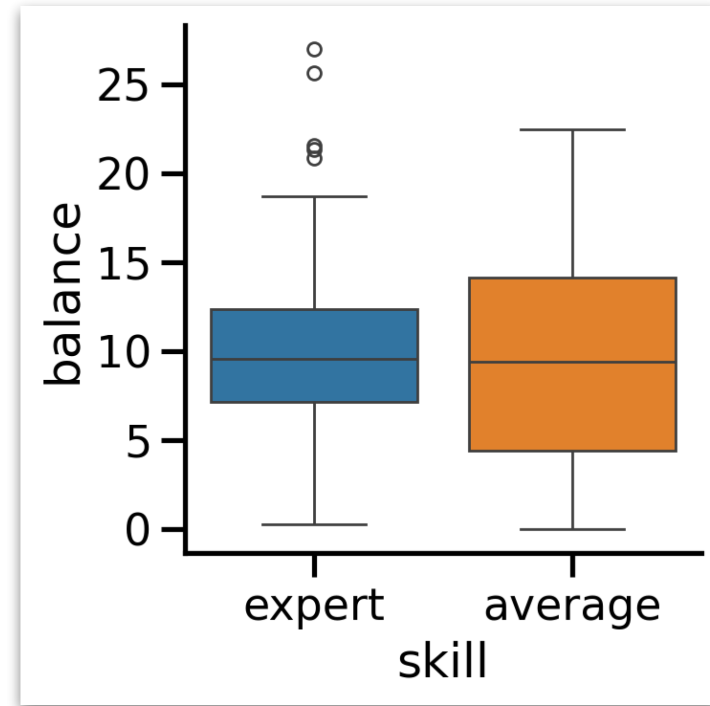


Do better hands affect balance differently based on skill?

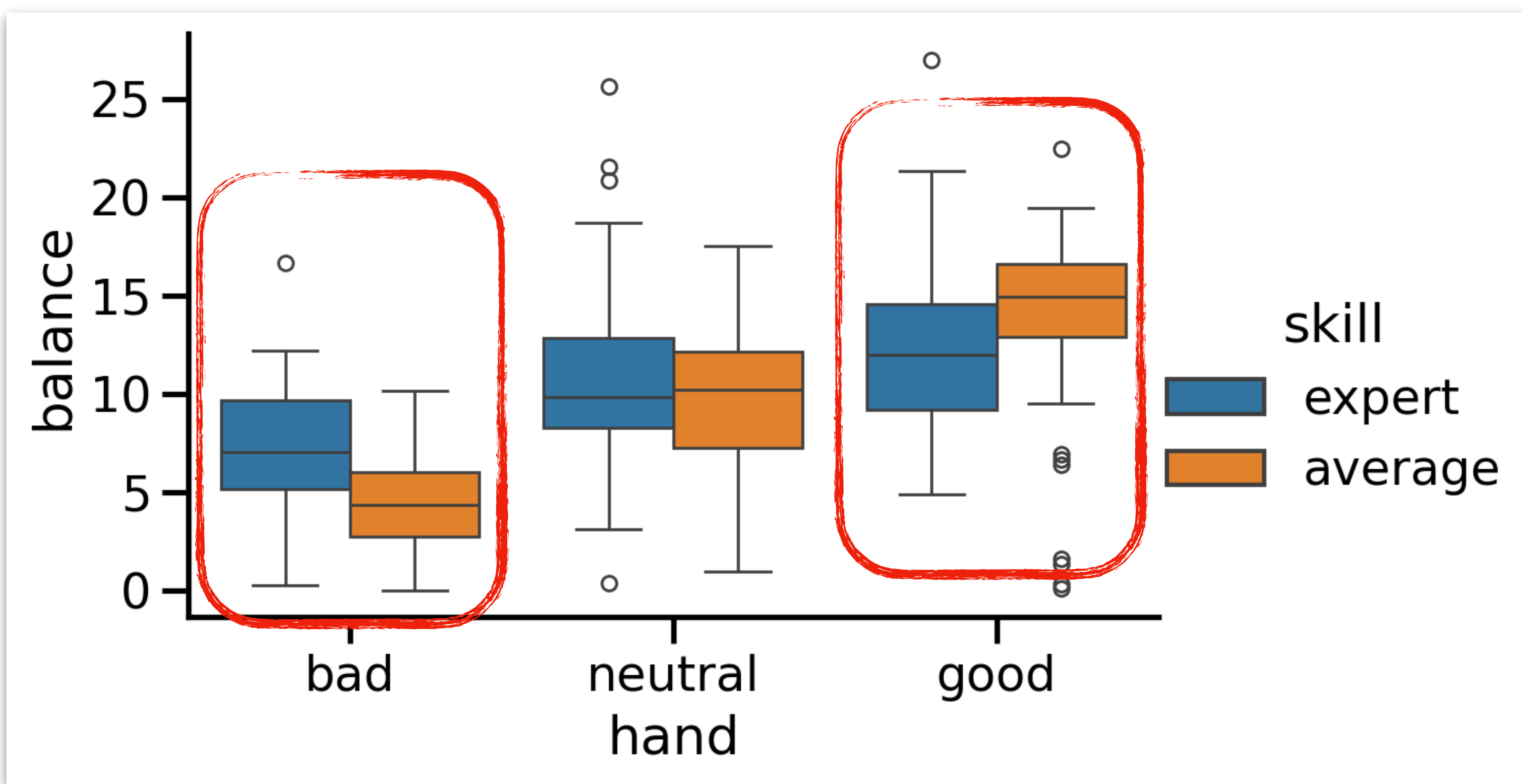
Effect of Hand



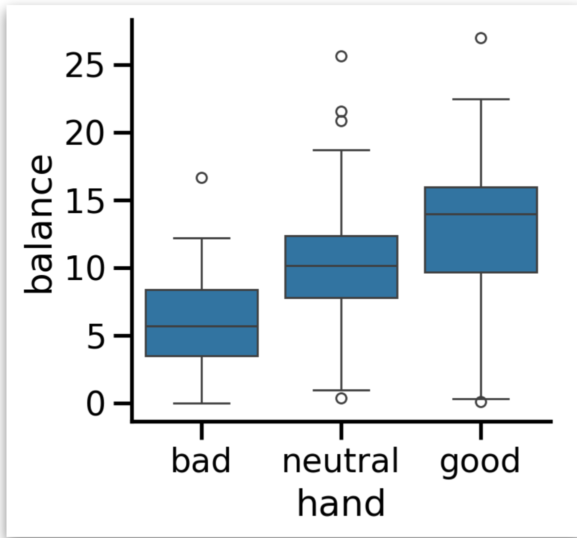
Effect of Skill

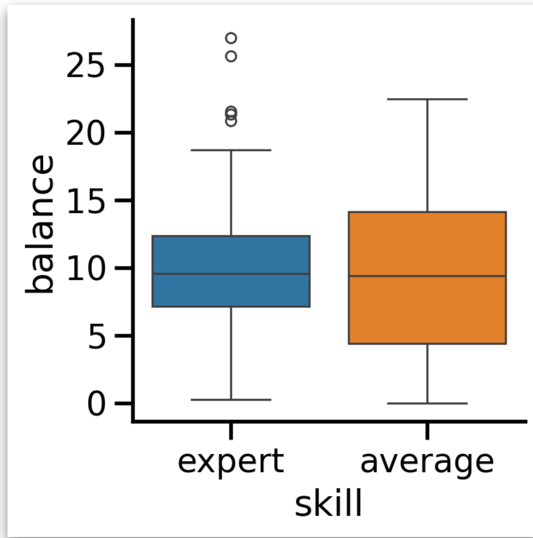


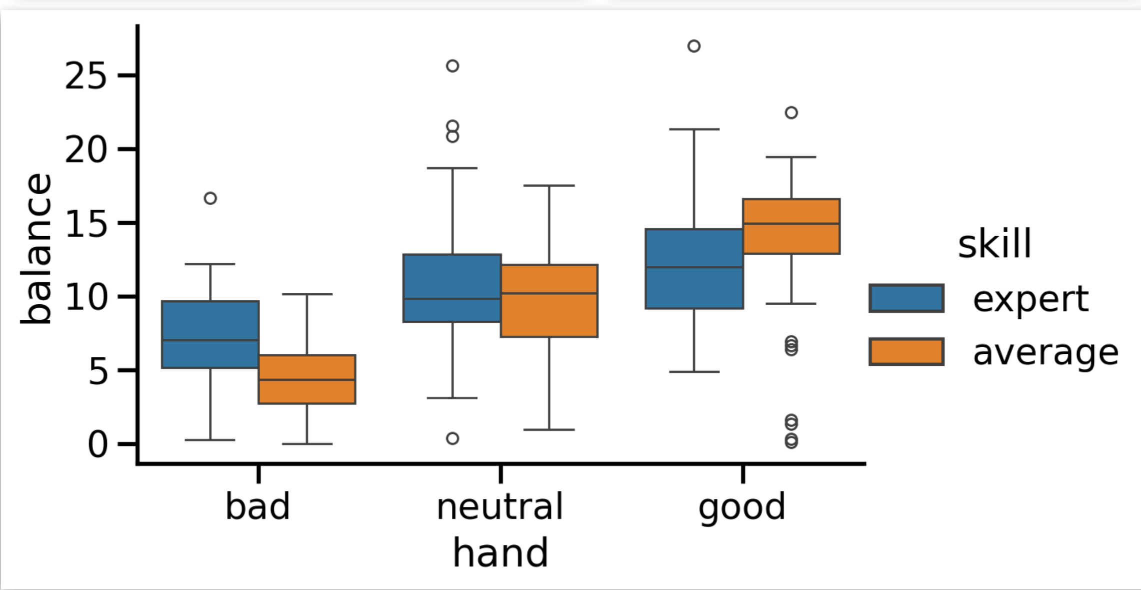
Hand & Skill



Let's estimate it…







Regression with **two categorical predictors**

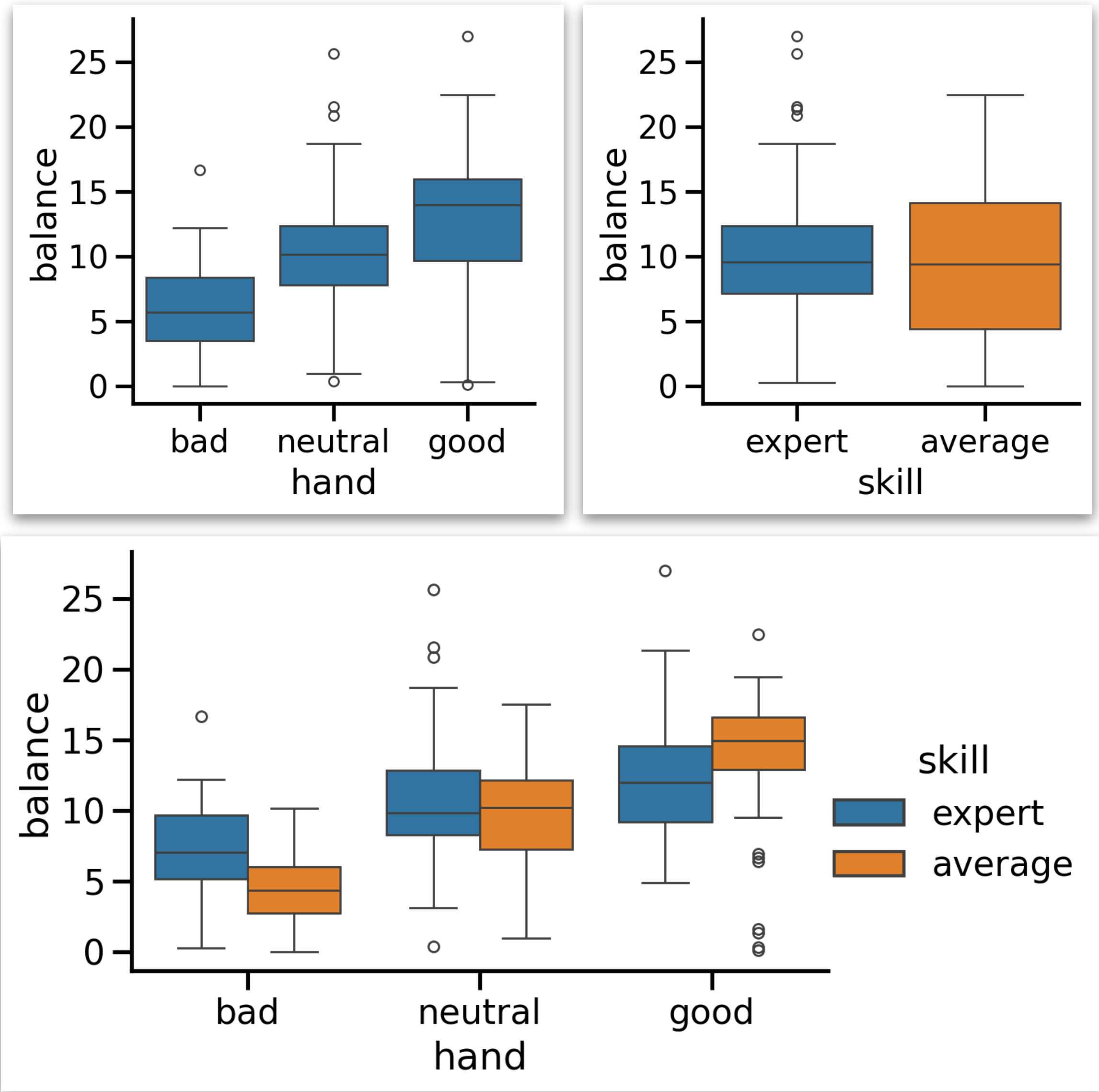
- $\text{balance}_i \sim \text{hand}_i * \text{skill}_i$

We need to **encode** hand (three levels) using (two) parameters

We need to **encode** skill (two levels) using (one) parameters

We need to **encode** the interaction (three levels) using (two) parameters

This a two-way Factorial ANOVA!



Regression with **two categorical predictors**

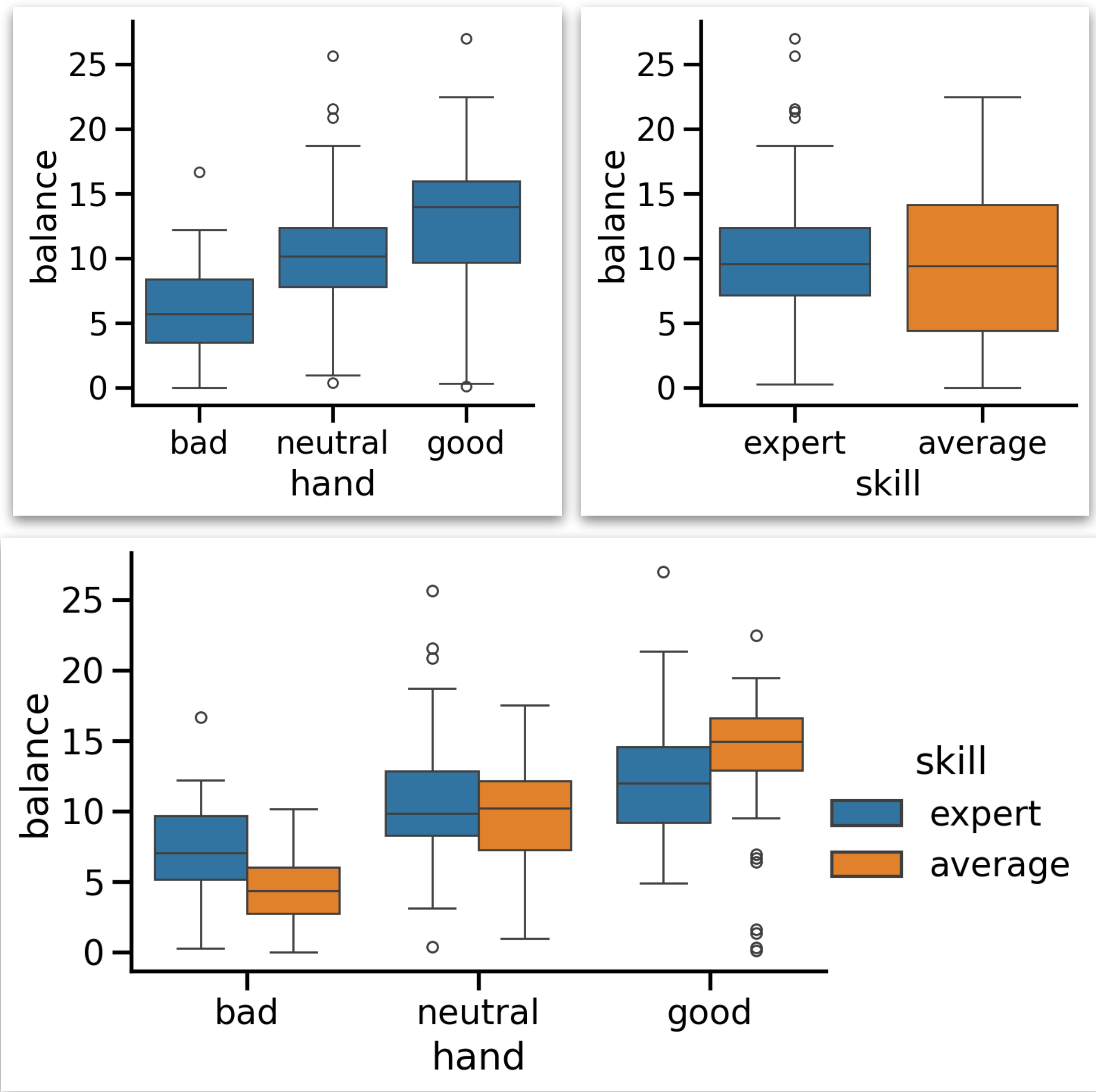
- $\text{balance}_i \sim \text{hand}_i * \text{skill}_i$

We need to **encode** hand (three levels) using (two) parameters

We need to **encode** skill (two levels) using (one) parameters

We need to **encode** the interaction (three levels) using (two) parameters

Two-way Factorial ANOVA



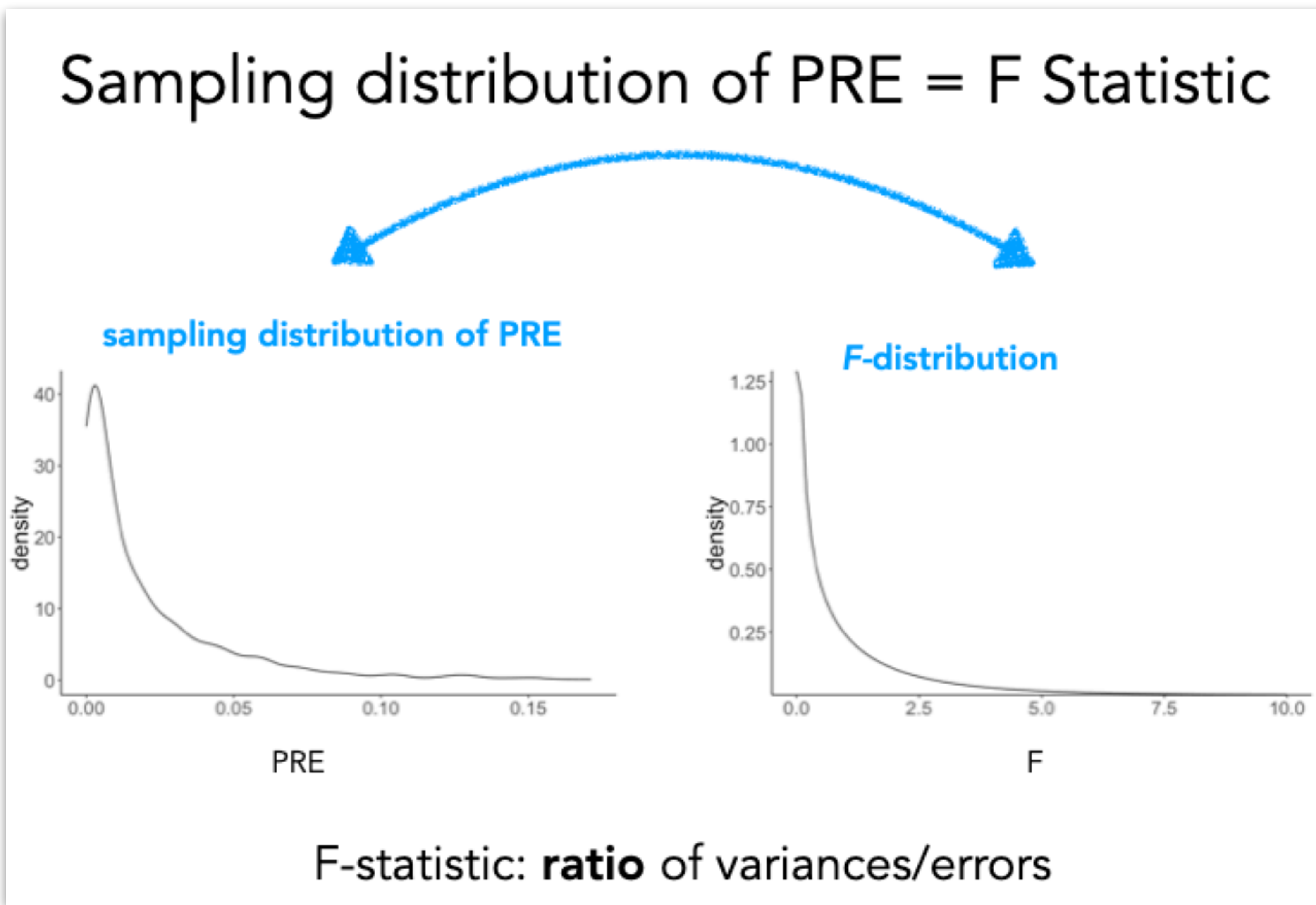
Factor 1: Hand

Factor 2: Skill

	bad	good	netural
expert			
average			

ANOVA: Analysis of Variance

- Regression with **only** categorical predictors ("factors")
- Developed by Ronald **F**isher (**F**-distribution!)



ANOVA: Analysis of Variance

- Regression with **only** categorical predictors ("factors")
- Developed by Ronald **F**isher (**F**-distribution!)
- Mathematical "trick" to calculate unique **variance** (model error) attributable to each **factor**

We care about *overall* variance across *all* levels

F is an **omnibus** or **joint** test

Factor 1: Hand

Factor 2: Skill

	bad	good	netural
expert			
average			

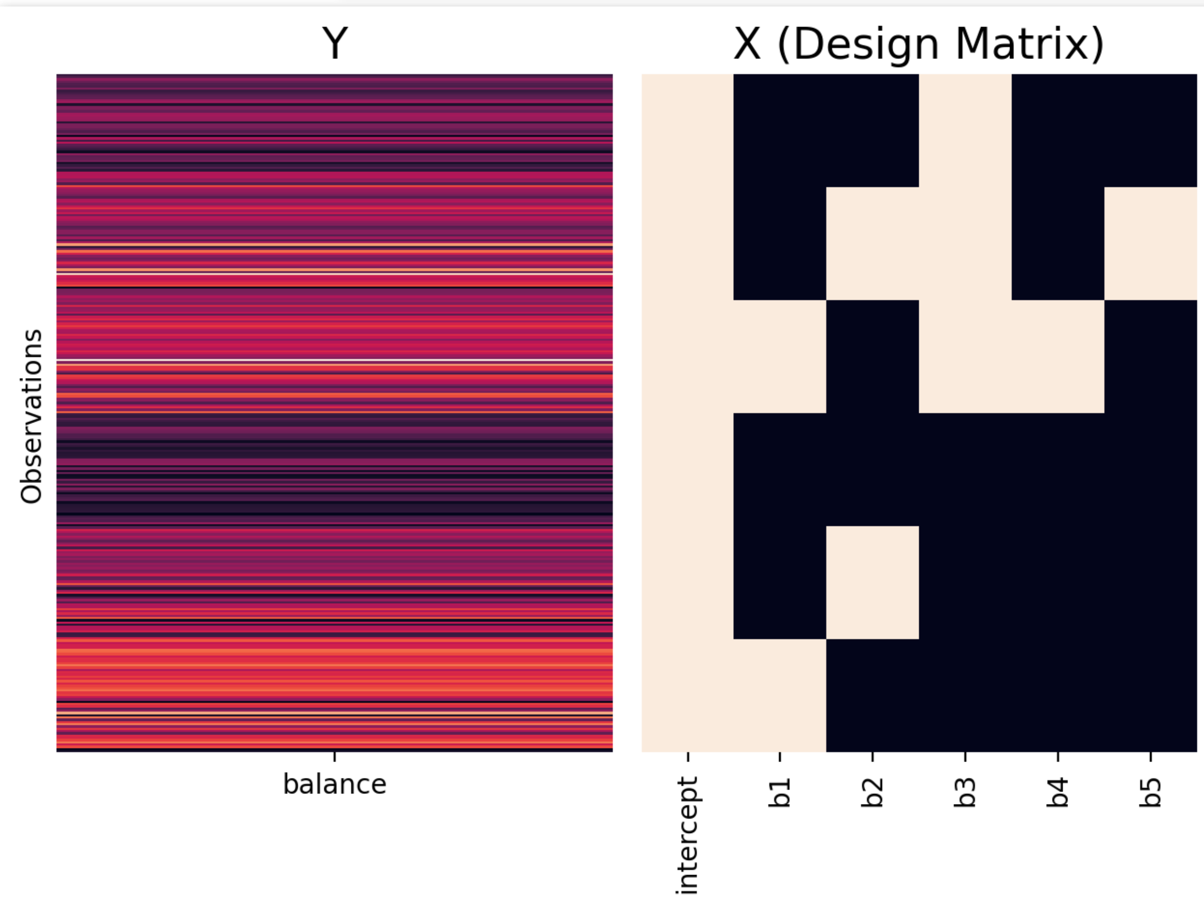
ANOVA: Analysis of Variance

- Regression with **only** categorical predictors ("factors")
- Developed by Ronald **F**isher (**F**-distribution!)
- Mathematical "trick" to calculate unique **variance** (model error) attributable to each **factor**
- **2 requirements**
  1. Use a valid **contrast** coding scheme
  2. Calculate variance using **type III** sum-of-squares approach

Valid contrast scheme(s)

- Codes should **sum-to-zero** across factor levels
- Comparisons should be **independent/orthogonal** across factor levels

```
twoway_treatment = ols('balance ~ C(hand) * C(skill)',
```




```
array([[1., 0., 0., 0., 0.],
       [0., 1., 0., 0., 0.],
       [0., 0., 1., 0., 0.],
       [0., 0., 0., 1., 0.],
       [0., 0., 0., 0., 1.]])
```

```
treatment_codes.sum(axis=0)
```

```
array([1., 1., 1., 1., 1.])
```

```
treatment_codes.sum(axis=1).sum()
```

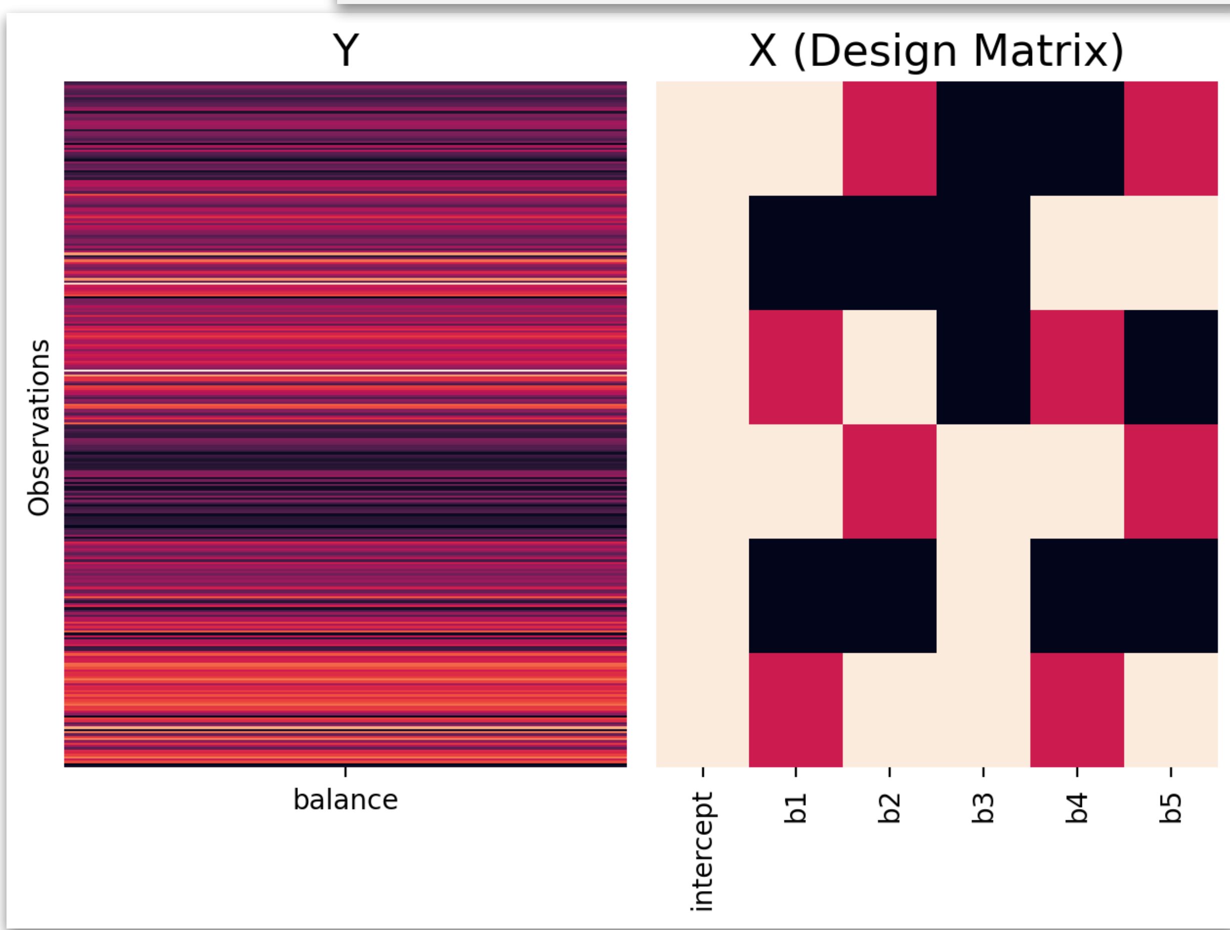
```
np.float64(5.0)
```

Treatment coding is not valid for ANOVA!

Valid contrast scheme(s)

- Codes should **sum-to-zero** across factor levels
- Comparisons should be **independent/orthogonal** across factor levels

```
twoway_sum = ols('balance ~ C(hand, Sum) * C(skill, Sum)',
```




```
array([[ 1.,  0.,  0.,  0.],
       [ 0.,  1.,  0.,  0.],
       [ 0.,  0.,  1.,  0.],
       [ 0.,  0.,  0.,  1.],
       [-1., -1., -1., -1.]])
```

```
sum_codes.sum(axis=0)
```

```
array([0., 0., 0., 0.])
```

```
sum_codes.sum(axis=1).sum()
```

```
np.float64(0.0)
```

Deviation (sum) coding is valid for ANOVA!

ANOVA: Analysis of Variance

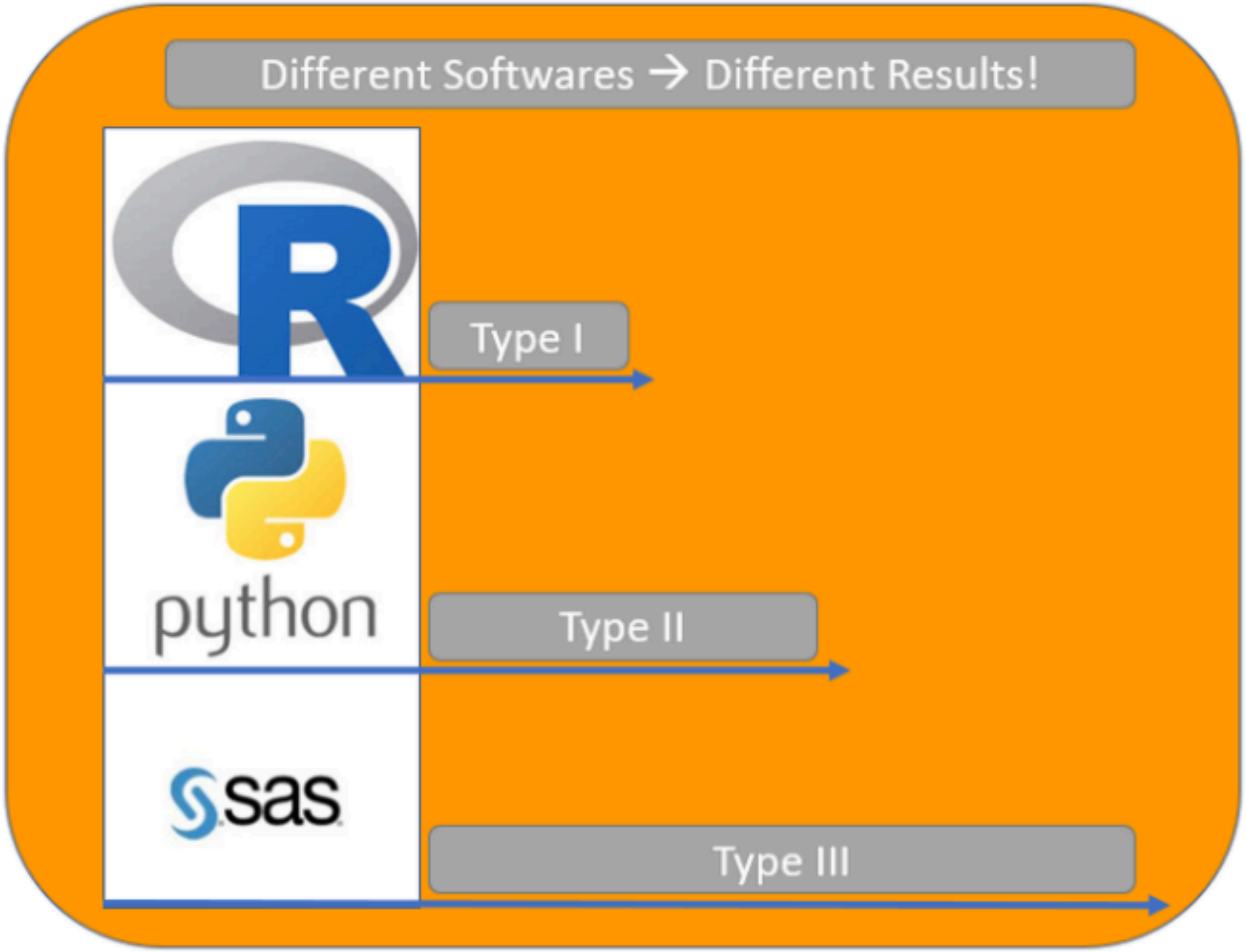
- Regression with **only** categorical predictors ("factors")
- Developed by Ronald **F**isher (**F**-distribution!)
- Mathematical "trick" to calculate unique **variance** (model error) attributable to each **factor**
- **2 requirements**
  1. Use a valid **contrast** coding scheme
  2. Calculate variance using **type III** sum-of-squares approach

The different sums of squares

Three different methodologies for splitting variation exist: Type I, Type II and Type III Sums of Squares. They do not give the same result in case of unbalanced data.

Type I, Type II and Type III ANOVA have different outcomes!

Default sums of squares ...



Default Types of Sums of Squares for different programming languages

not great for reproducibility ...

Type I Sums of Squares

Type I Sums of Squares are Sequential, so the order of variables in the models makes a difference. This is rarely what we want in practice!

Sums of Squares are Mathematically defined as:

- SS(A) for independent variable A
- SS(B | A) for independent variable B
- SS(AB | B, A) for the interaction effect

Type II Sums of Squares

Type II Sums of Squares should be used if there is no interaction between the independent variables.

Sums of Squares are Mathematically defined as:

- SS(A | B) for independent variable A
- SS(B | A) for independent variable B
- No interaction effect

solution: always use `anova_lm(model, typ=3)`

caution: this is what `anova_lm()` uses by default

Type III Sums of Squares

The Type III Sums of Squares are also called partial sums of squares again another way of computing Sums of Squares:

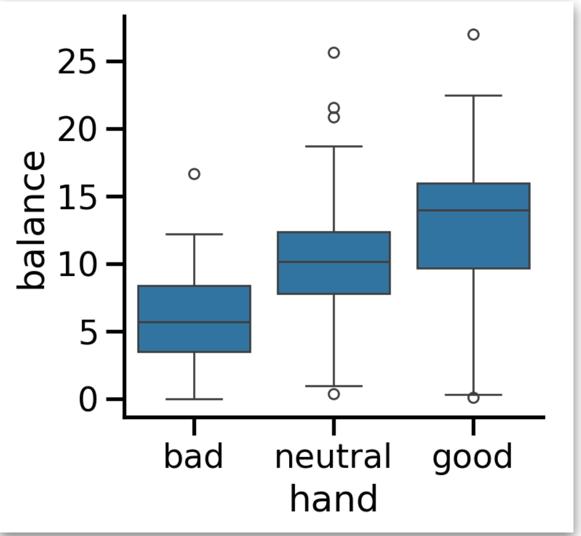
- Like Type II, the Type III Sums of Squares are not sequential, so the order of specification does not matter.
- Unlike Type II, the Type III Sums of Squares do specify an interaction effect.

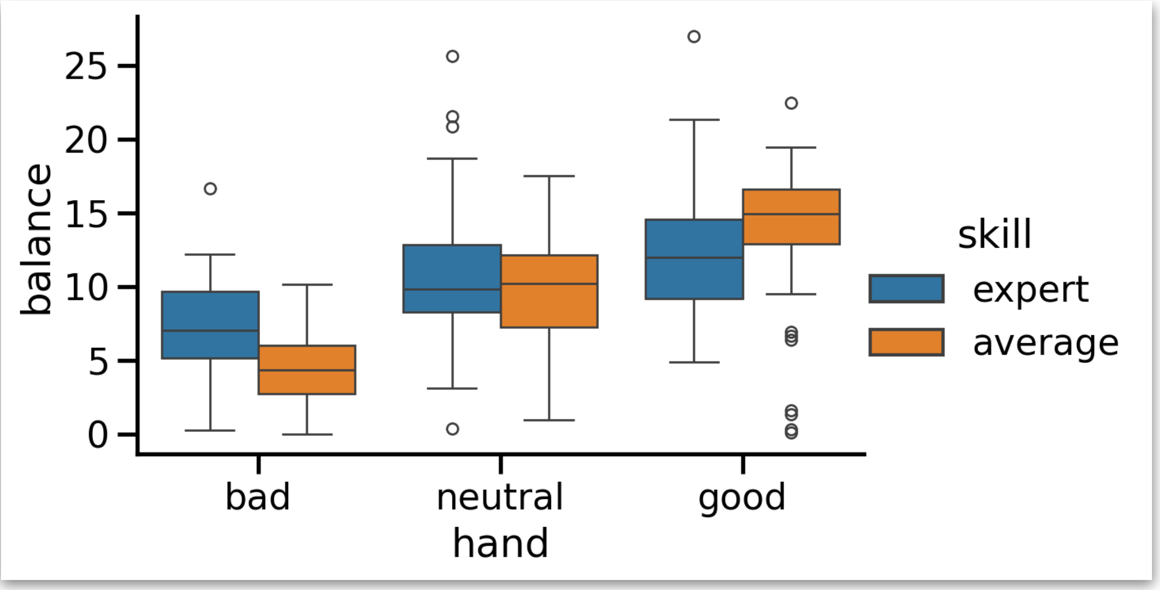
Sums of Squares are Mathematically defined as:

- SS(A | B, AB) for independent variable A
- SS(B | A, AB) for independent variable B

This is the standard in the literature

Treatment coding: Let's see what happens...





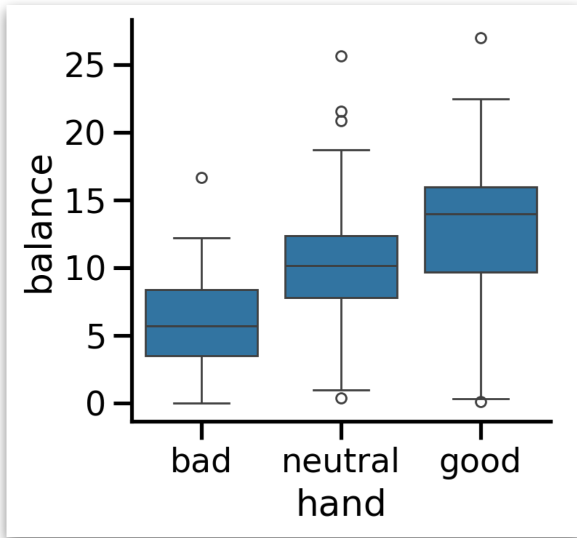

```
2  twoway_treatment = ols('balance ~ C(hand) * C(skill)',
3
4  anova_lm(twoway_treatment,typ=3).round(4)
```

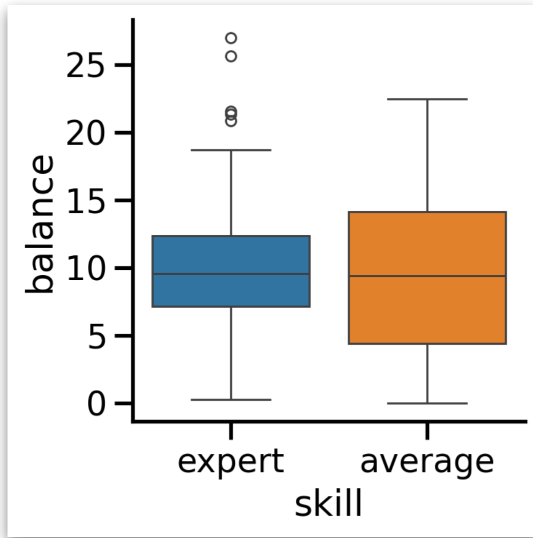
✓ 0.0s

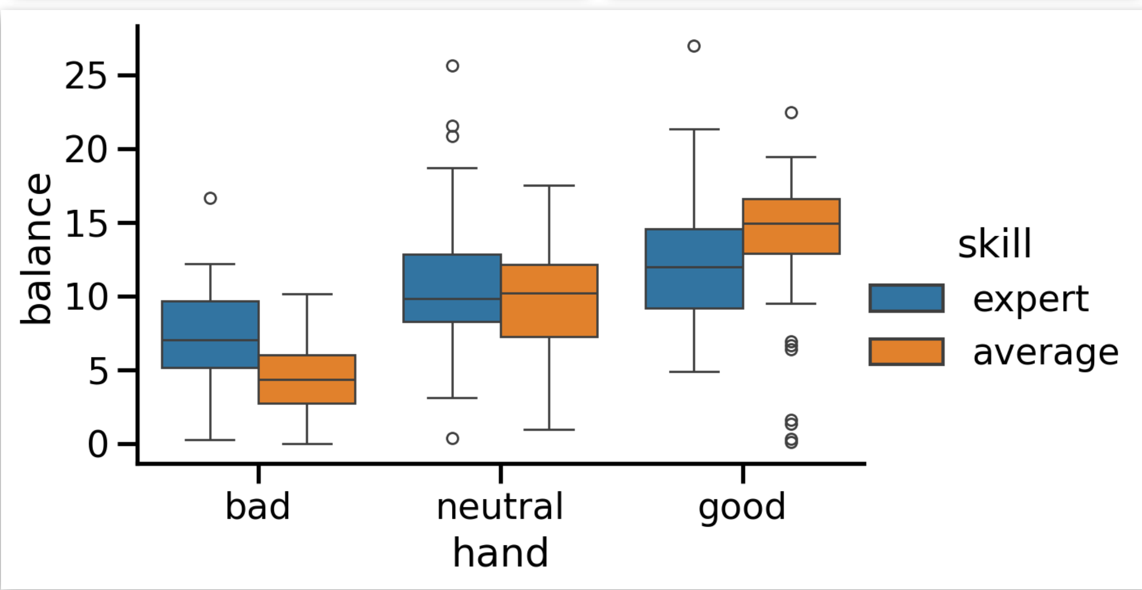
	sum_sq	df	F	PR(>F)
Intercept	1051.8450	1.0	65.0728	0.0000
C(hand)	2135.2201	2.0	66.0481	0.0000
C(skill)	183.5754	1.0	11.3570	0.0009
C(hand):C(skill)	228.9817	2.0	7.0830	0.0010
Residual	4752.2521	294.0	NaN	NaN

Ummm...

Sum coding: Let's see what happens…







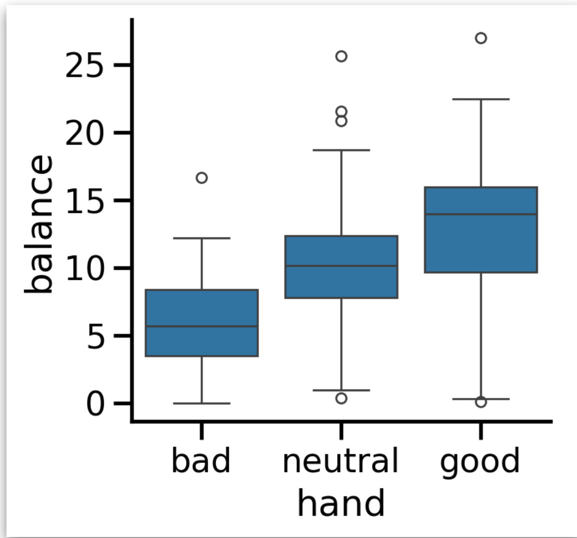

```
twoway_sums = ols('balance ~ C(hand,Sum) * C(skill, Sum)',

anova_lm(twoway_sums, typ=3).round(4)
```

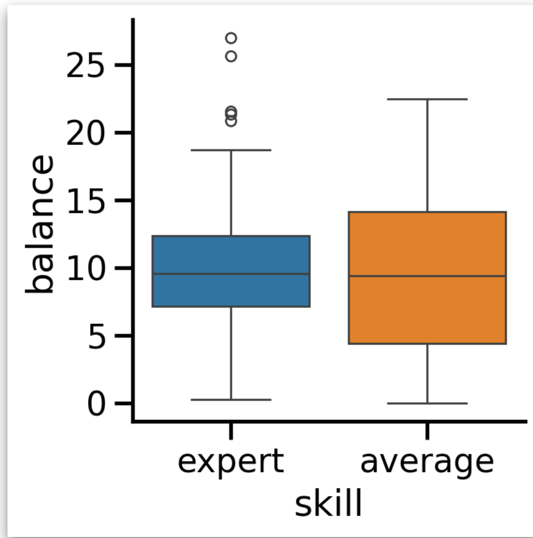
✓ 0.0s

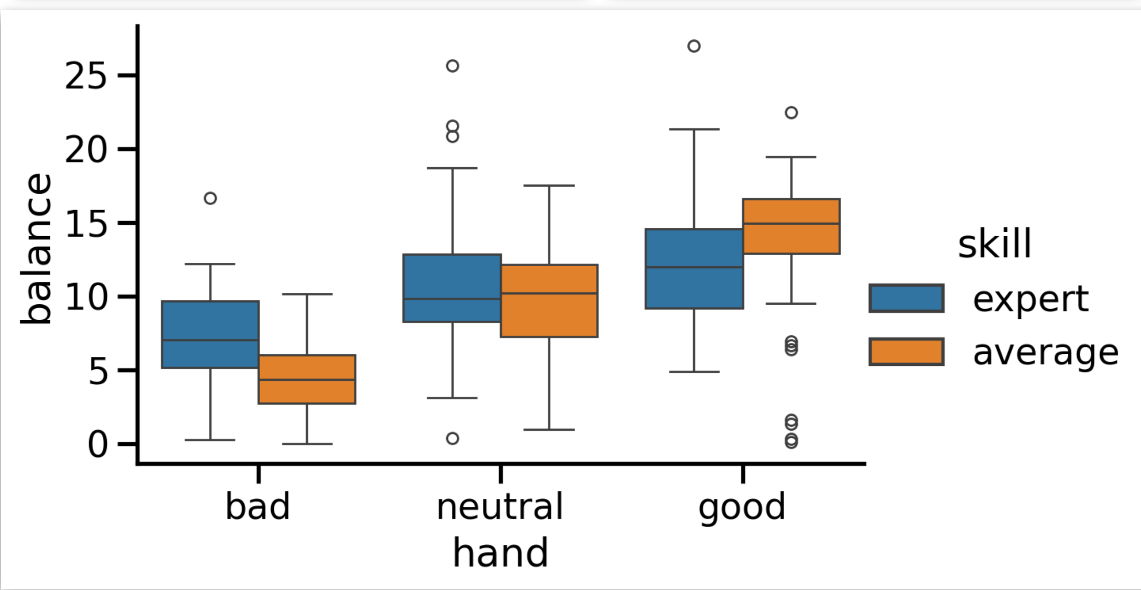
	sum_sq	df	F	PR(>F)
Intercept	28644.6637	1.0	1772.1137	0.0000
C(hand, Sum)	2559.4014	2.0	79.1692	0.0000
C(skill, Sum)	39.3494	1.0	2.4344	0.1198
C(hand, Sum):C(skill, Sum)	228.9817	2.0	7.0830	0.0010
Residual	4752.2521	294.0	NaN	NaN

Much better!

Valid: Two-way ANOVA








```python
twoway_sums = ols('balance ~ C(hand,Sum) * C(skill, Sum)',

anova_lm(twoway_sums, typ=3).round(4)
```

✓ 0.0s

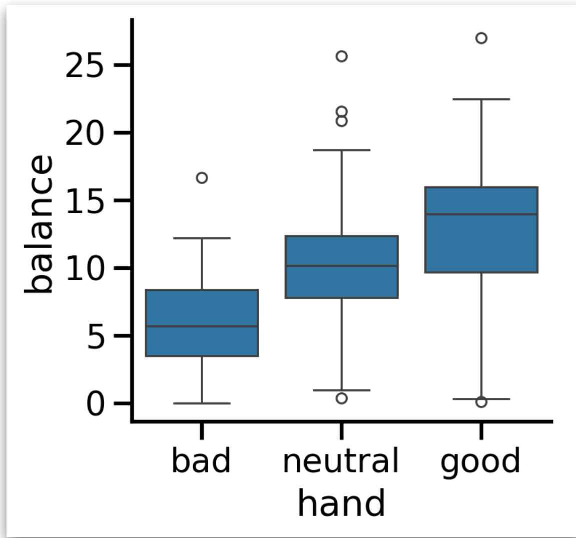
	sum_sq	df	F	PR(>F)
Intercept	28644.6637	1.0	1772.1137	0.0000
C(hand, Sum)	2559.4014	2.0	79.1692	0.0000
C(skill, Sum)	39.3494	1.0	2.4344	0.1198
C(hand, Sum):C(skill, Sum)	228.9817	2.0	7.0830	0.0010
Residual	4752.2521	294.0	NaN	NaN

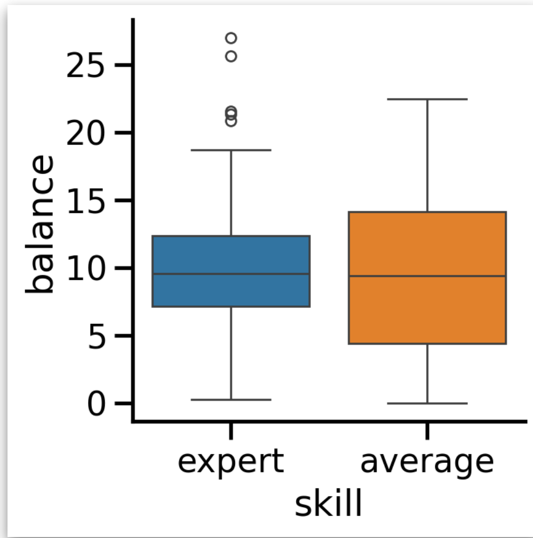
Main effect of **hand**

No main effect of skill

Interaction between hand and skill

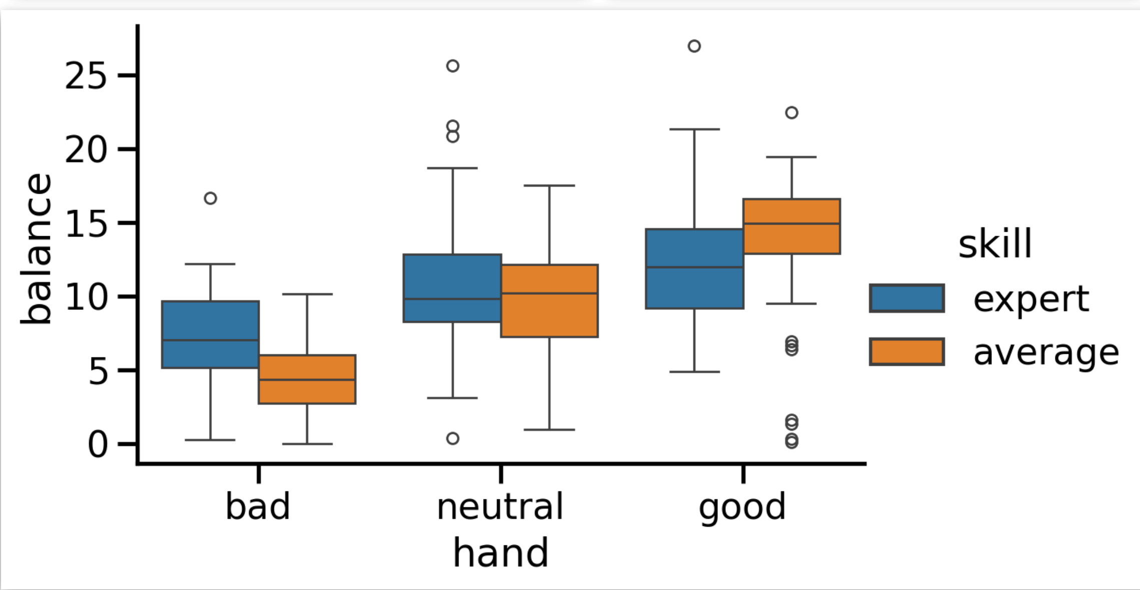
Reporting Two-way ANOVA





	sum_sq	df	F	PR(>F)
Intercept	28644.6637	1.0	1772.1137	0.0000
C(hand, Sum)	2559.4014	2.0	79.1692	0.0000
C(skill, Sum)	39.3494	1.0	2.4344	0.1198
C(hand, Sum):C(skill, Sum)	228.9817	2.0	7.0830	0.0010
Residual	4752.2521	294.0	NaN	NaN

There was no main effect of skill $F(1, 294) = 2.43$, $p = .12$. The final balance of average ($M = 9.41$, $SD = 5.51$) and expert poker players ($M = 10.13$, $SD = 4.50$) did not differ significantly.



The quality of a player's hand significantly affected the final balance $F(2, 294) = 79.17$, $p < .001$. The final balance for good hands (M = 13.03, SD = 4.65) was significantly greater than for neutral hands (M = 10.35, SD = 4.24), and the balance for neutral hands was significantly higher than for bad hands (M = 5.94, SD = 3.34).

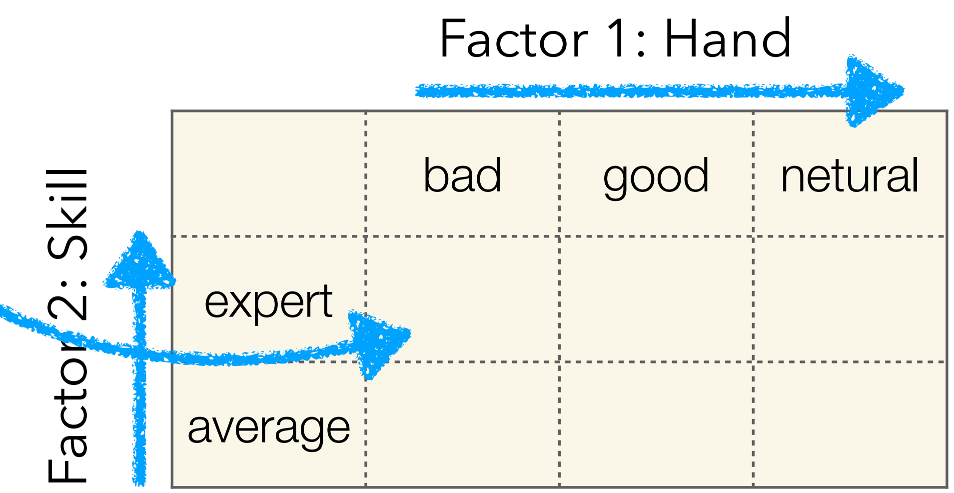
However, this was *moderated* by a significant hand x skill interaction $F(2, 294) = 7.08$, $p < .001$. Such that for bad hands, average players had a lower final balance than experts, for good hands, average players had a higher final balance than experts.

ANOVA Summary

- ANOVA = "categorical multiple regression"
- ANCOVA = "categorical & continuous multiple regression"
  - We didn't discuss this today but….HW 3….
- When? You care about **overall effect** of factor not specific level comparisons (we'll revisit…)
  - **Omnibus (joint)** F-tests over factor levels accounting for other factors
  - Can be **significant** even if individual parameter estimates are not!
- Requires **consideration** of how you *encode* categorical predictors
  - Does **not** affect overall model fit
  - Changes what parameter estimates mean
- Requires **consideration** of how you analyze unique variance
  - Type I, II, or III sums of squares

ANOVA Tips

unequal observations per cell



- Coding schemes matter when
  - 2+ predictors
  - 3+ factor levels
  - Unbalanced designs
- Always use **type III** sums of squares
  - Works properly for both balanced & **unbalanced** data
  - Unaffected by the *order* in which you put predictors into model
- **Always** use a valid **contrast** coding scheme
  - Stick with “Sum” or “Poly”

Next time…

- Interpreting parameter estimates…
- Making **additional comparisons** beyond *k*-1…
  - Parameterizing models for **planned comparisons**
  - Performing **post-hoc comparisons**

Break

On your own

- Update last week's github repo (`wk6-lab-yourID`)
- You should see
  - `06_models_solutions` (check your work today)
  - `07_models_solutions` (check your work today)
  - `08_models` (work through today)

Make sure you commit and push your work before you leave!