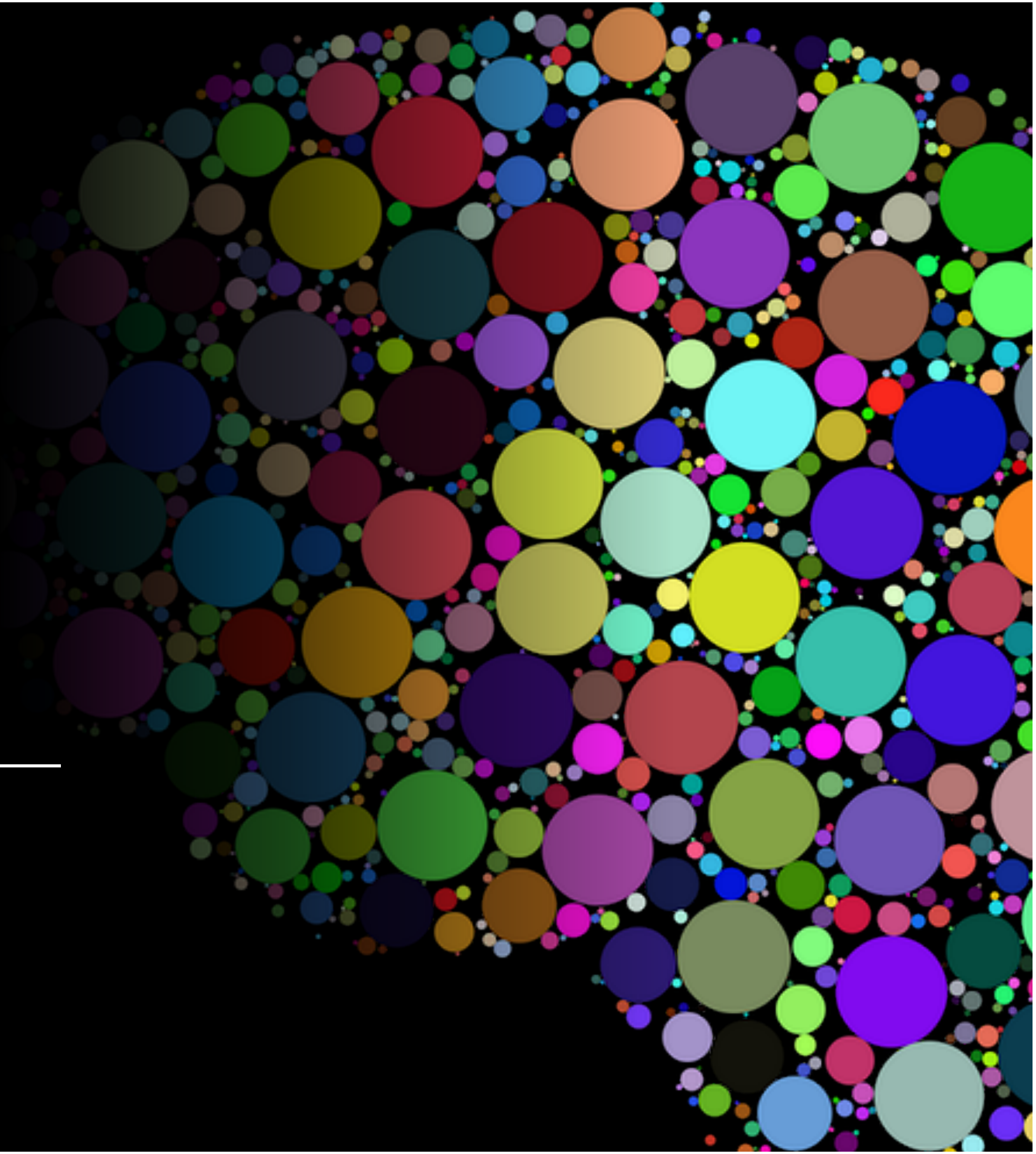
Between-Groups ANOVA

Chapter 12

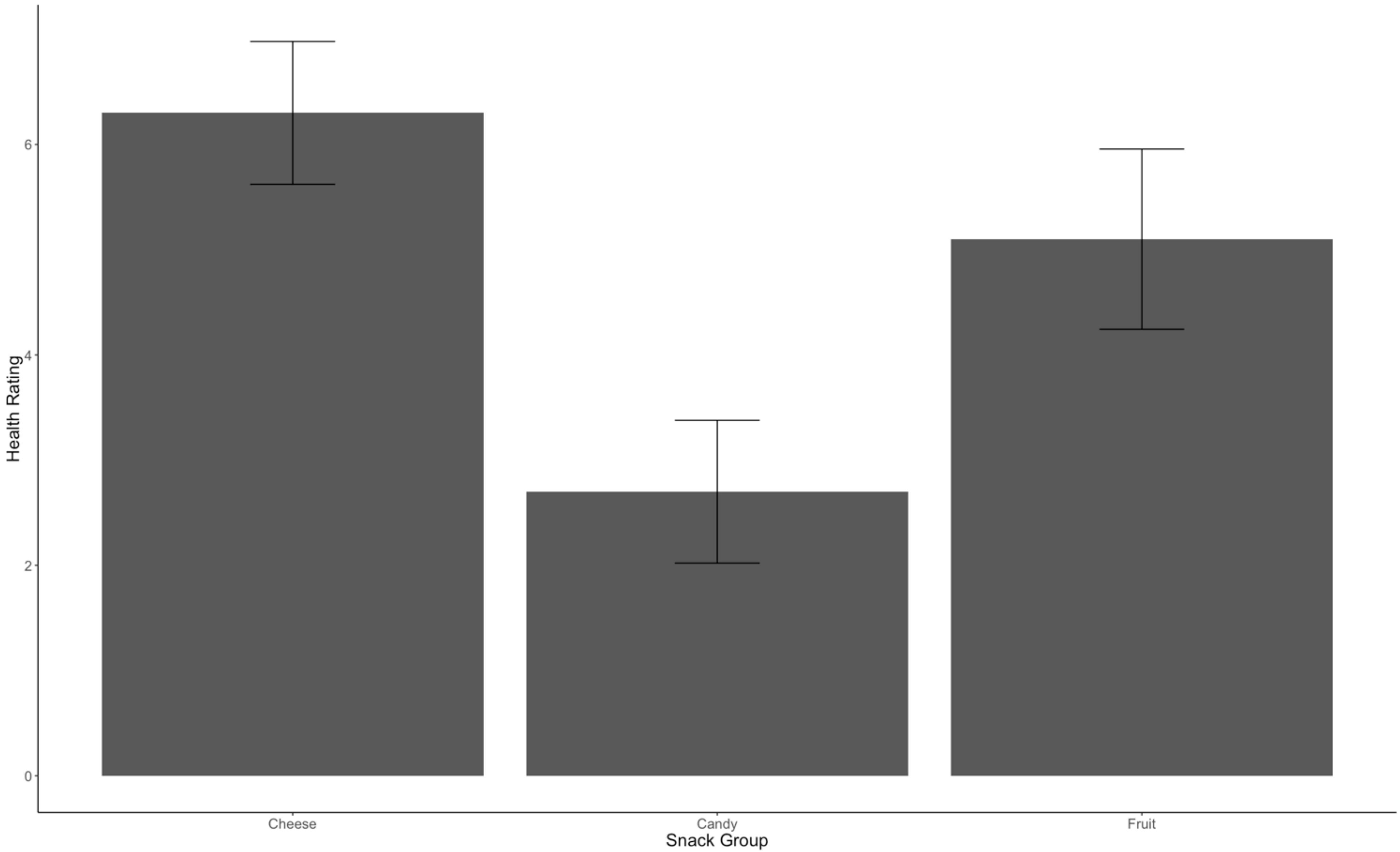


Quick Test Reminder

- One person = Z score
- One sample with population standard deviation = Z test
- One sample no population standard deviation = single t-test
- One sample test twice = paired samples t
- Two samples = independent t

Quick Distributions Reminder

- Z = Distribution of scores
- Z = distribution of means (for samples)
- t = distribution of means (for samples with estimated standard deviation)
- t = distribution of mean differences (for paired samples with estimated standard deviation)
- t = distribution of differences between means (for two groups independent t)



Health Rating
6
4
2
0
Cheese
Candy
Fruit
Snack Group

So what now?

- We could do lots of *pairwise* *t*-tests
  - Group 1 versus Group 2
  - Group 1 versus Group 3
  - Group 2 versus Group 3
- (that’s too easy though)

Why not use multiple *t*-tests?

- The problem of too many *t* tests
  - Fishing for a finding
- Problem of Type I error (alpha)
  - New type 1 error rate = $1 - (1\text{-alpha})^c$
  - We want type 1 error rate to stay at .05

TABLE 12-1. The Probability of a Type I Error Increases as the Number of Statistical Comparisons Increases

As the number of samples increases, the number of *t* tests necessary to compare every possible pair of means increases at an even greater rate. And with that, the probability of a Type I error quickly becomes far larger than 0.05.

Number of Means	Number of Comparisons	Probability of a Type I Error
2	1	0.05
3	3	0.143
4	6	0.265
5	10	0.401
6	15	0.537
7	21	0.659

So what do we do?

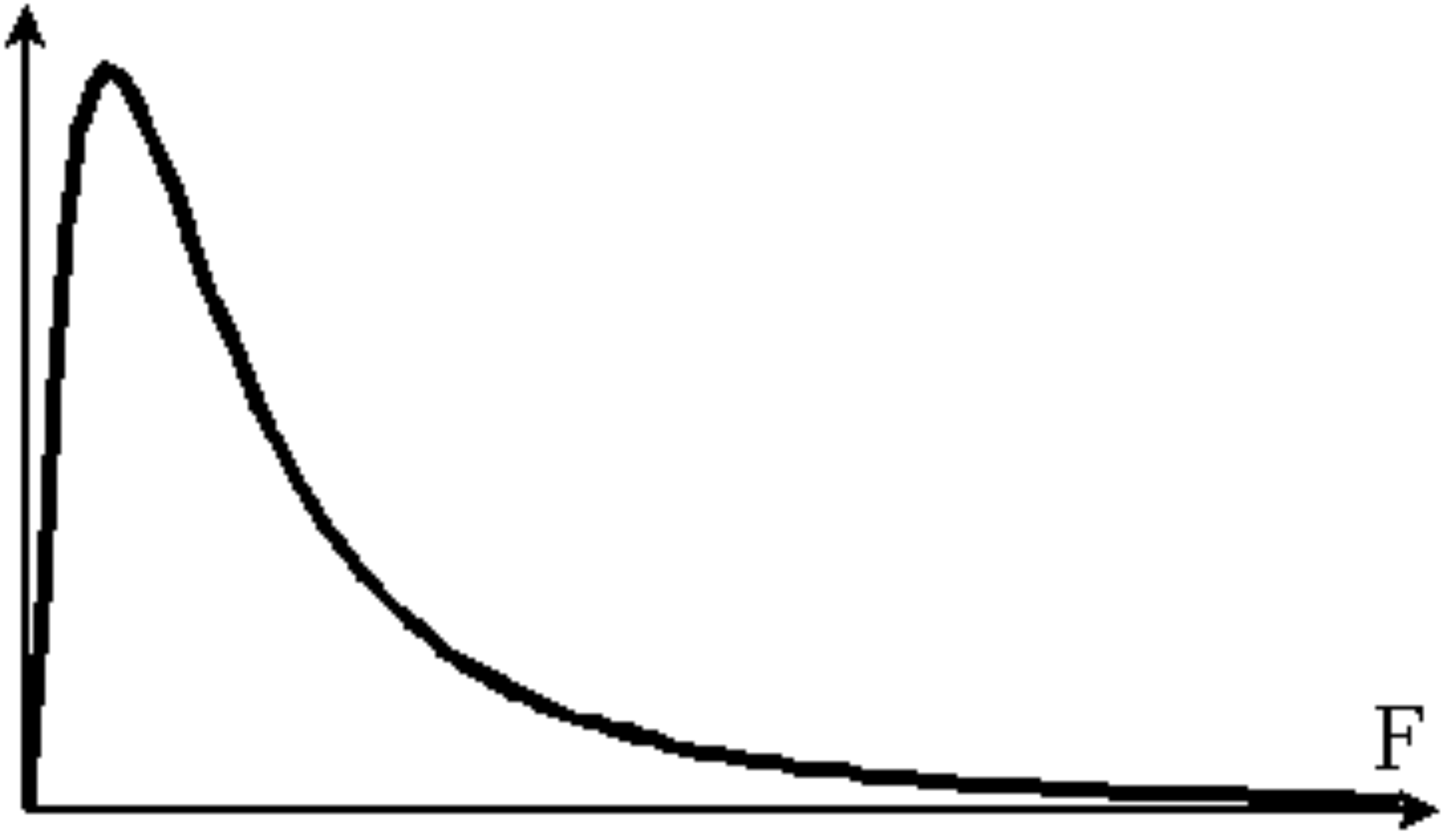
- When to use an *F* distribution
  - Working with more than two samples
- ANOVA
  - Analysis of Variance
  - Used with two or more nominal independent variables and an interval/ratio dependent variable

Types of ANOVA

> One-Way: hypothesis test including one nominal variable with more than two levels and a scale DV

- Within-Groups: more than two samples, with the same participants; also called repeated-measures
- Between-Groups: more than two samples, with different participants in each sample

The F distribution



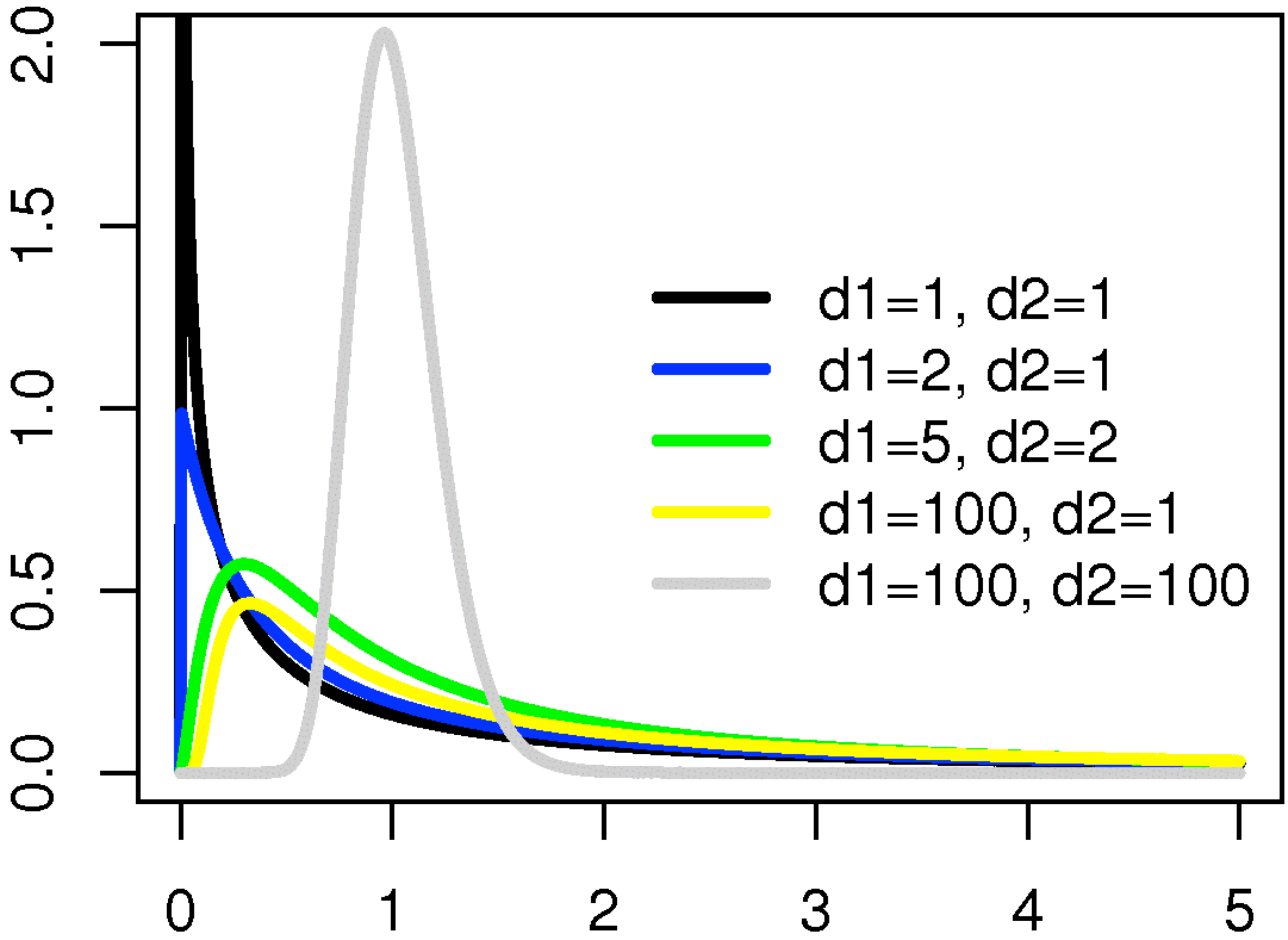

d1=1, d2=1
d1=2, d2=1
d1=5, d2=2
d1=100, d2=1
d1=100, d2=100
0.0
0.5
1.0
1.5
2.0
0
1
2
3
4
5

TABLE 12-2. Connections Among Distributions

The z distribution is subsumed under the t distributions in certain specific circumstances, and both the z and t distributions are subsumed under the F distributions in certain specific circumstances.

	When Used	Links Among the Distributions
z	One sample; μ and σ are known	Subsumed under the t and F distributions
t	(1) One sample; only μ is known (2) Two samples	Same as z distribution if there is a sample size of ∞ (or just very large)
F	Three or more samples (but can be used with two samples)	Square of z distribution if there are only two samples and a sample size of ∞ (or just very large); square of t distribution if there are only two samples

Let's talk about test statistics

Test type	Formula
Z	$\frac{M - \mu_M}{\sigma_M}$
Single t	$\frac{M - \mu_M}{s_M}$
Paired t	$\frac{M}{s_M}$
Independent t	$\frac{M - M}{s_{difference}}$

Now what do we do with 3 or more Ms and SDs?

The F Distribution

> Analyzing variability to compare means

- $F = \dfrac{\text{variance between groups}}{\text{variance within groups}}$

> That is, the difference among the sample means divided by the average of the sample variances

Logic behind the *F* Statistic

> Quantifies overlap

> Two ways to estimate population variance

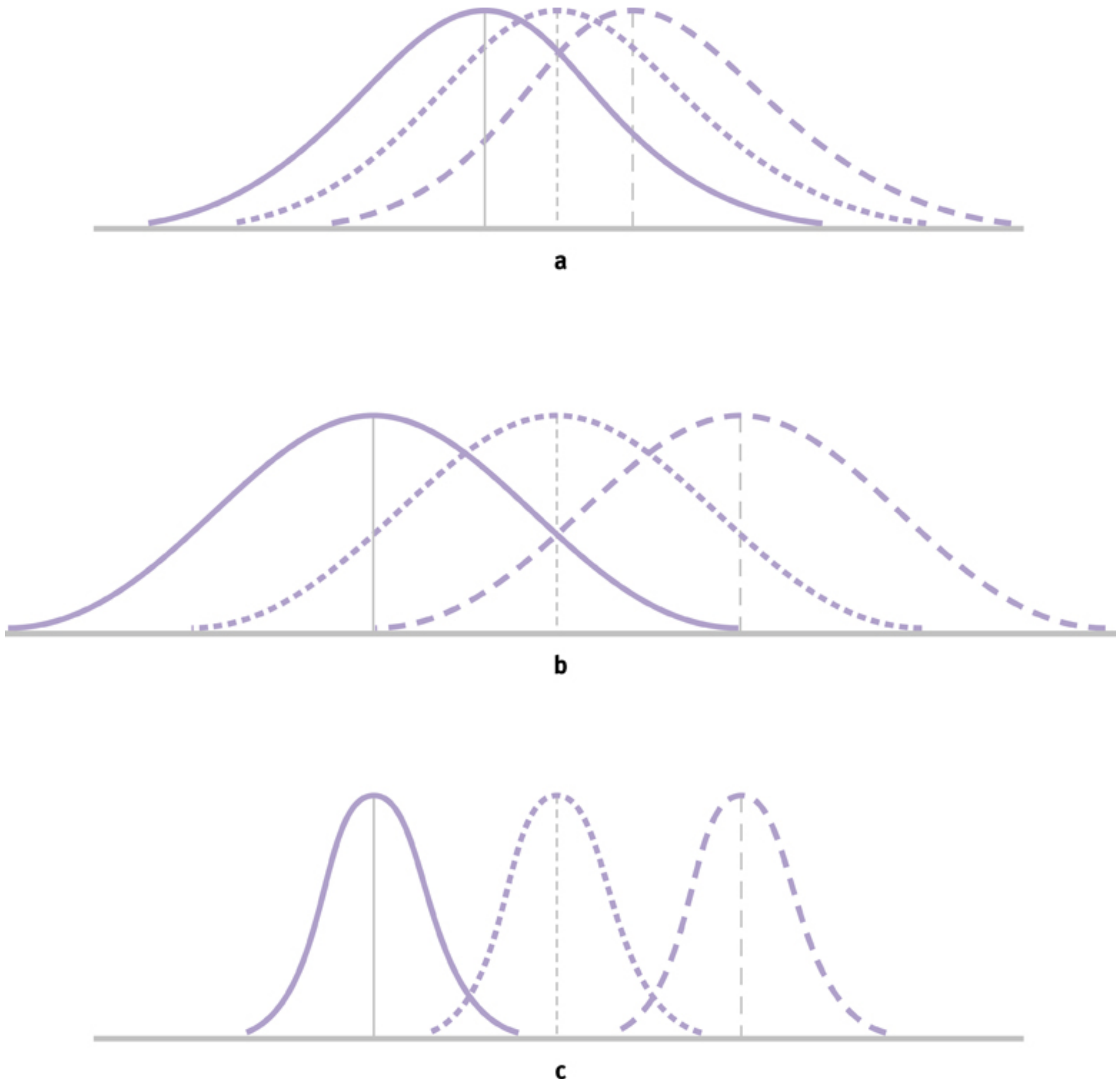
- Between-groups variability
- Within-groups variability

Types of Variance

> Between groups: estimate of the population variance based on differences among group means

> Within groups: estimate of population variance based on differences within (3 or more) sample distributions

The Logic of ANOVA



Formulae

$$SS_{total} = SS_{within} + SS_{between}$$

$$SS_{total} = \sum (X - GM)^2$$

$$SS_{within} = \sum (X - M)^2 \qquad MS_{within} = \frac{SS_{within}}{df_{within}} \qquad F = \frac{MS_{between}}{MS_{within}}$$

$$SS_{between} = \Sigma(M - GM)^2 n \qquad MS_{between} = \frac{SS_{between}}{df_{between}}$$

The Source Table

> Presents important calculations and final results in a consistent, easy-to-read format

TABLE 12-4. The Source Table Organizes Our ANOVA Calculations

A source table helps researchers organize the most important calculations necessary to conduct an ANOVA as well as the final results. The numbers 1–5 in the first row are used in this particular table only to help you understand the format of source tables; they would not be included in an actual source table.

1 Source	2 *SS*	3 *df*	4 *MS*	5 *F*
Between	$SS_{between}$	$df_{between}$	$MS_{between}$	F
Within	SS_{within}	df_{within}	MS_{within}	
Total	SS_{total}	df_{total}		

TABLE 12-9. A Source Table with Formulas

This table summarizes the formulas for calculating an *F* statistic.

Source	*SS*	*df*	*MS*	*F*
Between	$\Sigma(M - GM)^2 * n$	$N_{groups} - 1$	$\frac{SS_{between}}{df_{between}}$	$\frac{MS_{between}}{MS_{within}}$
Within	$\Sigma(X - M)^2$	$df_1 + df_2 + \ldots + df_{last}$	$\frac{SS_{within}}{df_{within}}$	
Total	$\Sigma(X - GM)^2$	$N_{total} - 1$		

[Expanded formula: $df_{within} = (N_1 - 1) + (N_2 - 1) + \ldots + (N_{last} - 1)$]

Sum of Squares Example

- Let’s try calculating SS within and SS between.
  - Load the chapter 12 data in R.

Sum of Squares Example

- Let’s try SS total
  - Each person minus grand mean, squared, and summed.
- The logic here is overall variance, both due to your IV and error will be calculated.

Sum of Squares Example

- mean(unlist(chapter12))
- sum((chapter12 – 4.7)^2)

- 96.3

Sum of Squares Example

- SS within:
  - The difference of each person minus their group mean, squared.
  - That's the top half of the variance equation.
- The logic here is that we don't know why people differ within their own group, so it's considered error.

Sum of Squares Example

- summary(chapter12)
- sum((chapter12$cheese – 6.3)^2)
- sum((chapter12$candy – 2.7)^2)
- sum((chapter12$fruit – 5.10)^2)

- 8.1+8.1+12.9 = 29.1 within SS

Sum of Squares Example

> SS between
  - Each group minus the grand mean, squared, times n, and totaled.
  - Why times n? Every other formula is for each person, so we need to do this one for each person as well.

> The logic here is to measure how much of the overall variance is IV group differences, not individual differences.

Sum of Squares Example

```
> (6.3 - 4.7)^2 *10 +
> (2.7 - 4.7)^2 *10 +
> (5.1 - 4.7)^2 *10

> 67.2
```

Variance Type	SS	df	MS	F
Between groups (IV)	67.2			
Within Groups (error)	29.1			
Total	96.3			

Let’s talk about *df*

Test type	***df***
Single sample	N – 1
Paired samples t	N – 1
Independent t	N – 1 + N – 1

Degrees of Freedom for ANOVA

$$df_{between} = N_{groups} - 1$$

$$df_{within} = df_1 + df_2 + df_3 + \ldots df_{last}$$

$$df_1 = n_1 - 1$$

Variance Type	SS	df	MS	F
Between groups (IV)	67.2	2		
Within Groups (error)	29.1	27		
Total	96.3	29		

Variance Type	SS	df	MS	F
Between groups (IV)	67.2	2	DIVIDE! 33.6	
Within Groups (error)	29.1	27	1.08	
Total	96.3	29	XX	

Variance Type	SS	df	MS	F
Between groups (IV)	67.2	2	33.6	DIVIDE! 31.11
Within Groups (error)	29.1	27	1.08	XX
Total	96.3	29	XX	XX

One-Way Between-Groups ANOVA

- Everything about ANOVA but the calculations
  - 1. Identify the populations, distribution, and assumptions.
  - 2. State the null and research hypotheses.
  - 3. Determine the characteristics of the comparison distribution.
  - 4. Determine the critical value, or cutoff.
  - 5. Calculate the test statistic.
  - 6. Make a decision.

Assumptions of ANOVAs

> Random selection of samples
> Normally distributed sample
> DV is scale
> Homoscedasticity: samples come from populations with the same variance
  - Generally this is referred to as homogeneity

Step 2

> Label all the sample groups

> ONLY us = and =/

> Why?

Step 2

- R: cheese =/ candy =/ fruit
- N: cheese = candy = fruit

Step 3

- Here we want to calculate the means, so we know what we are comparing.
- Then calculate the ANOVA summary table.

Step 3

> To get the means:

- summary(*dataset*)

> To get the ANOVA table:

- Restructure the data
- Add a participant number
- Run ezANOVA

Step 3

- Restructure the data from wide to long format.
  - Use the reshape library (NOT reshape2)
- longdata = melt(*dataset,*
  measured = c("*column*", *column*"…)

Step 3

> Add a participant number:

> longdata$partno = 1:nrow(longdata)

Step 3

> NOTE:

- Reshape changes the names of the columns
- Variable = group, your IV
- Value = score, your DV

Filter

	variable	value	partno
1	cheese	6	1
2	cheese	7	2
3	cheese	8	3
4	cheese	6	4
5	cheese	5	5
6	cheese	6	6
7	cheese	7	7

Step 3

> Run ezANOVA to get your ANOVA.

ezANOVA(data = *dataset*,

dv = value,

between = variable,

wid = partno,

type = 3,

return_aov = T)

```
$ANOVA
    Effect DFn DFd        F            p p<.05       ges
2 variable   2  27 31.17526 9.629428e-08     * 0.6978193

$`Levene's Test for Homogeneity of Variance`
  DFn DFd       SSn  SSd         F         p p<.05
1   2  27 0.2666667 13.1 0.2748092 0.7618165

$aov
Call:
   aov(formula = formula(aov_formula), data = data)

Terms:
                variable Residuals
Sum of Squares      67.2      29.1
Deg. of Freedom        2        27

Residual standard error: 1.038161
Estimated effects may be unbalanced
```

Step 4

> Calculate the F critical?

- qf(.05, df bn, df wn, lower.tail = F)
- qf(.05, 2, 27, lower.tail = F)

- Don’t change lower.tail
- No / 2 since always one tailed

Step 5

> Use F from the source table.

Step 6

- Is F found in step 5 > F critical in step 4?
  - Reject null
- If not:
  - Fail to reject null

Calculating Effect Size

> R^2 is a common measure of effect size for ANOVAs.

$$R^2 = \frac{SS_{between}}{SS_{total}}$$

Amount of variance due to the IV out of the total.

> 67.2 / 96.3 = .70

> Or use the *ges* column.

Only 0-1, it is a proportion.

TABLE 12-12. Cohen's Conventions for Effect Sizes: R^2

The following guidelines, called *conventions* by statisticians, are meant to help researchers decide how important an effect is. These numbers are not cutoffs, merely rough guidelines to aid researchers in their interpretation of results.

Effect Size	Convention
Small	0.01
Medium	0.06
Large	0.14

Post-Hoc Tests to Determine Which Groups Are Different

> When you have three groups, and *F* is significant, how do you know where the difference(s) are?

> Run a post hoc test

> Run a post hoc correction

- Bonferroni
- (other types: Tukey, Scheffe)

Important Programming Note

- Post hoc *tests*
  - Independent t because it matches the type of research design (between subjects ANOVA = between groups = independent t)
- Post hoc *corrections*
  - Solve our type 1 error problem

The Bonferroni Test

- A post-hoc correction on a test that provides a more strict critical value for every comparison of means.
- We use a smaller critical region to make it more difficult to reject the null hypothesis.
  - Determine the number of comparisons we plan to make.
- Divide the *alpha* ($p<.05$, $p<.01$) level by the number of comparisons.

TABLE 12-14. The Bonferroni Test: Few Groups, Many Comparisons

Even with a few means, we must make many comparisons to account for every possible difference. Because we run the risk of incorrectly rejecting the null hypothesis just by chance if we run so many tests, it is a wise idea to use a more conservative procedure, such as the Bonferroni test, when comparing means. The Bonferroni test requires that we divide an overall p level, such as 0.05, by the number of comparisons we will make.

Number of Means	Number of Comparisons	Bonferroni p Level (overall $p = 0.05$)
2	1	0.05
3	3	0.017
4	6	0.008
5	10	0.005
6	15	0.003
7	21	0.002

Bonferroni in *R*

```
> pairwise.t.test(longdata$value,
      longdata$variable,
      paired = F,
      var.equal = T,
      p.adjust.method = "bonferroni")
```

Bonferroni in *R*

```
        Pairwise comparisons using t tests with pooled SD

data:  longdata$value and longdata$variable

      cheese  candy
candy 7.3e-08 -
fruit 0.046   5.8e-05

P value adjustment method: bonferroni
```

Remember, you can turn off scientific notation with options(scipen = 999)

So, what do I do with this?

```
        Pairwise comparisons using t tests with pooled SD

data:  longdata$value and longdata$variable

      cheese      candy
candy 0.000000073 -
fruit 0.046       0.000057987

P value adjustment method: bonferroni
> 
```

Group 1	Group 2	P-value	Comparison	Reject?
Cheese **M = 6.3**	Candy M = 2.7			
Cheese **M = 6.3**	Fruit M = 5.1			
Candy **M = 2.7**	Fruit M = 5.1			

First, fill in the group comparisons

Group 1	Group 2	P-value	Comparison	Reject?
Cheese **M = 6.3**	Candy M = 2.7	< .001		
Cheese **M = 6.3**	Fruit M = 5.1	.046		
Candy **M = 2.7**	Fruit M = 5.1	< .001		

Then fill in the p values from the output

Group 1	Group 2	P-value	Comparison	Reject?
Cheese **M = 6.3**	Candy M = 2.7	< .001	$p < .05$	
Cheese **M = 6.3**	Fruit M = 5.1	.046	$p < .05$	
Candy **M = 2.7**	Fruit M = 5.1	< .001	$p < .05$	

Fill in your alpha value, remember p value is corrected already

Group 1	Group 2	P-value	Comparison	Reject?
Cheese M = 6.3	Candy M = 2.7	< .001	$p < .05$	YES Cheese > candy
Cheese M = 6.3	Fruit M = 5.1	.046	$p < .05$	YES Cheese > fruit
Candy M = 2.7	Fruit M = 5.1	< .001	$p < .05$	YES Fruit > candy

Decide if you want to reject or not!
Interpret!

Interpret!

> Therefore,

- Cheese > fruit > candy on ratings of healthiness for snacks.
- Because cheese is the best. ☺