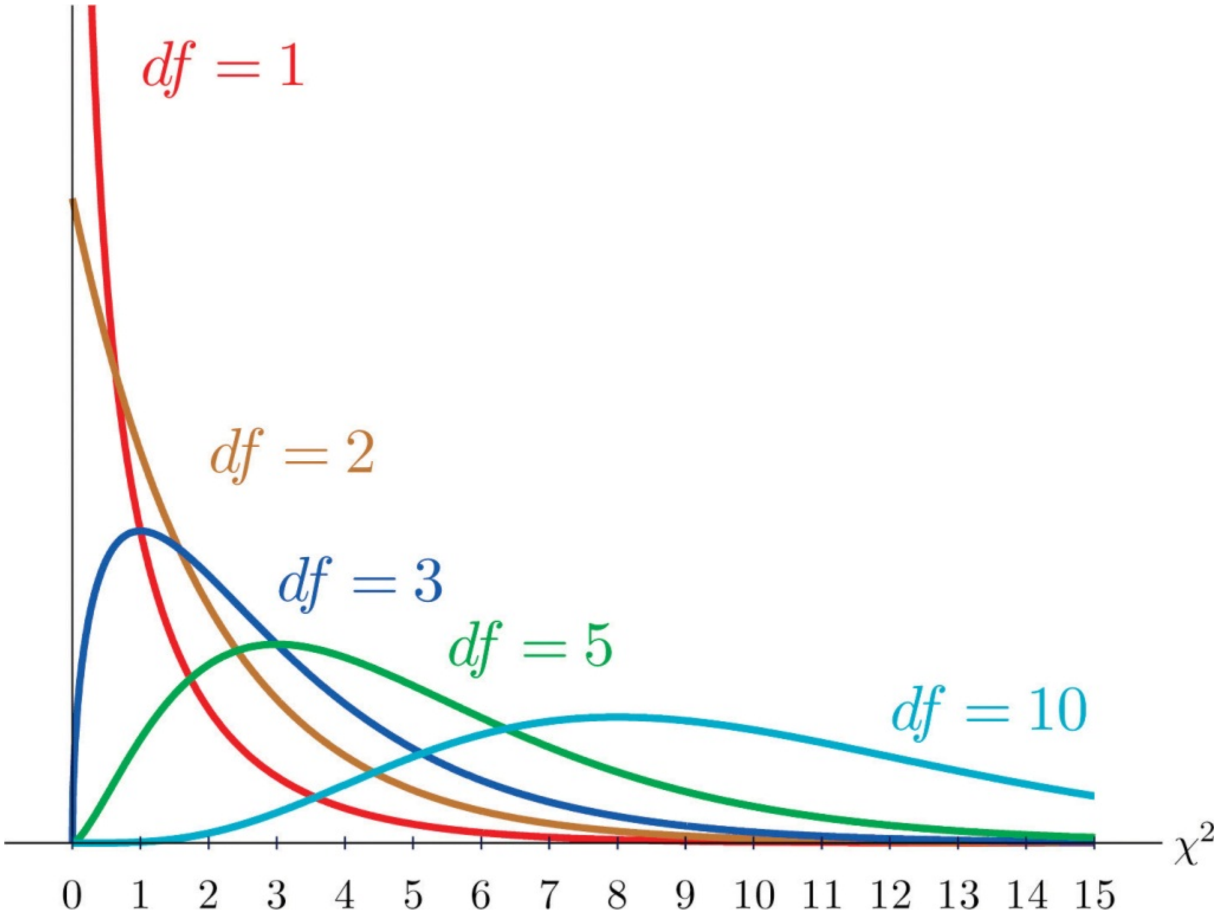
Chi square and F-test

Chi square:
Test independence of two random variables or follow a specific distribution

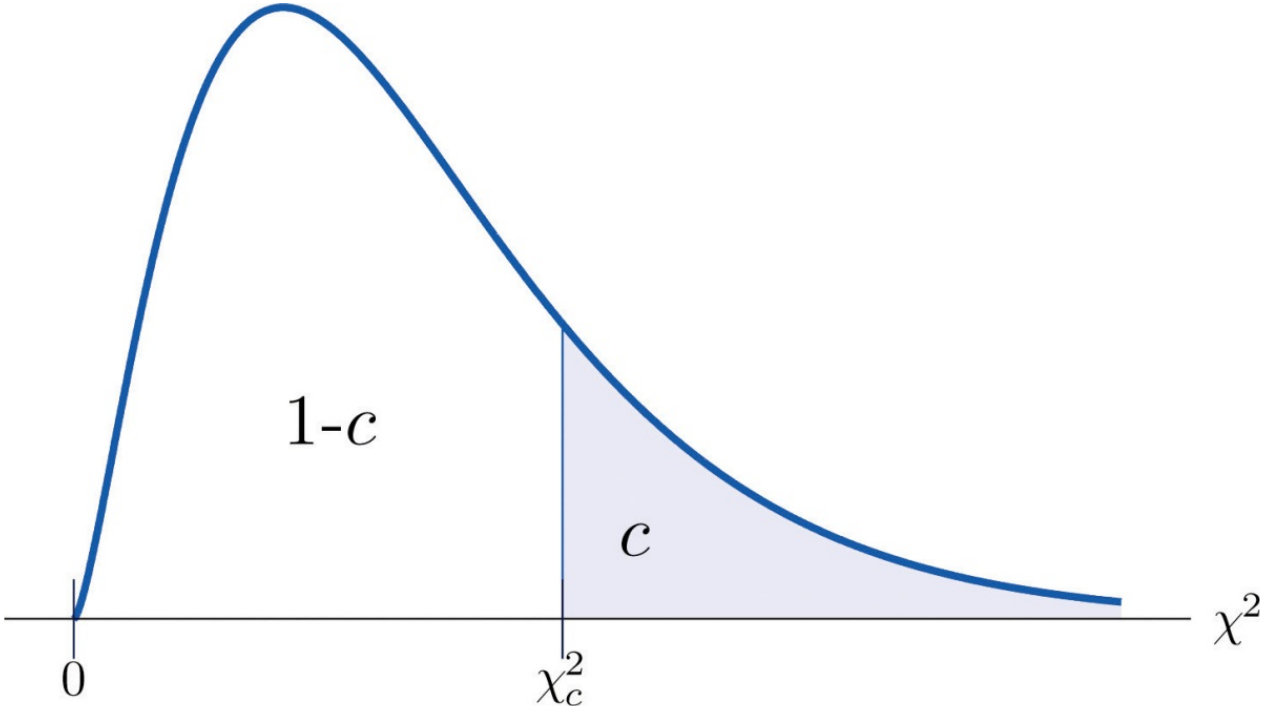
F-test:
Test if 2 population variances are equal
Or three or more population means are equal

Chi distribution



- Similar to the t-distribution, the chi distribution has a df parameter

Critical value for Chi distribution



- In analogy to the t-test, we estimate the critical value

Are these two samples independent

- We analyze if two random variables X1 and X2 to take their values independently $P(X1 \cap X2)=P(X1)P(X2)$

  Does baby girls have higher heart rate than male babies?

  Formulated as a test of the following hypotheses:

  H0 : Baby gender and baby heart rate are independent

  vs. Ha : Baby gender and baby heart rate are not independent

Gender and heart rate example

- Force separation of BPM into high and low heart rate.

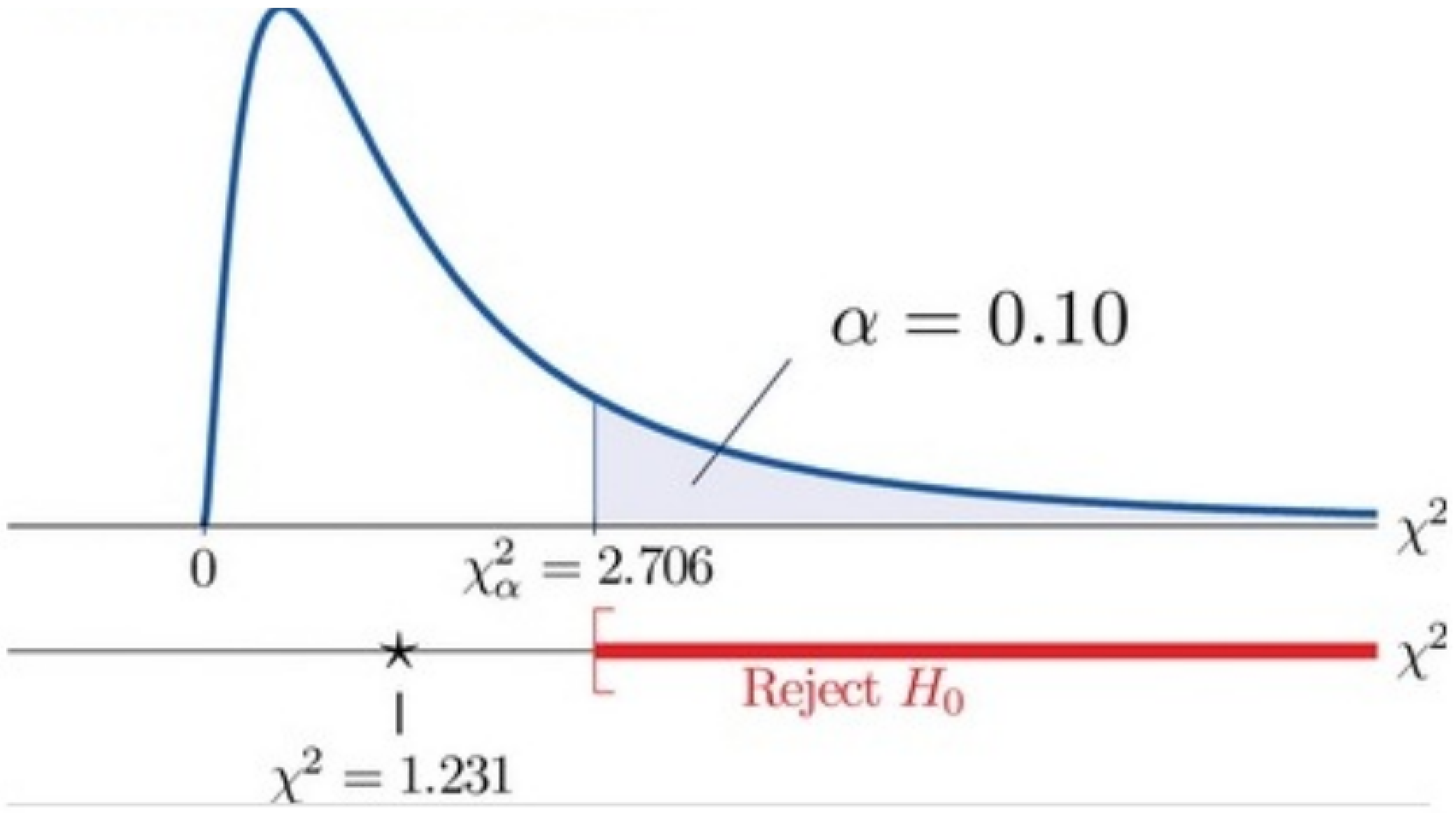
		Heart Rate		
		Low	**High**	**Row Total**
Gender	Girl	**11**	**7**	18
	Boy	**17**	**5**	22
Column Total		28	12	Total = 40

- Estimate expected values, assuming independence: Row total x col total / N_TOTAL
- With df = (C-1)*(R-1) = 1

		Heart Rate		
		Low	High	Row Total
Gender	Girl	$O=11 E=12.6$	$O=7 E=5.4$	R = 18
	Boy	$O=17 E=15.4$	$O=5 E=6.6$	R = 22
Column Total		C = 28	C = 12	n = 40

- Calculate the sum of square differences: Σ(O−E)^2 / E

$$\Sigma \frac{(O-E)^2}{E} = \frac{(11-12.6)^2}{12.6} + \frac{(7-5.4)^2}{5.4} + \frac{(17-15.4)^2}{15.4} + \frac{(5-6.6)^2}{6.6} = 1.231$$



- H0 cannot be rejected
- Notice this time we need to set qchisq(0.1,df=1,lower.tail = F)

How to assume a distribution with chi square test ?

- We consider die. Is the die fair/uniform distributed ? We take n=60 samples.

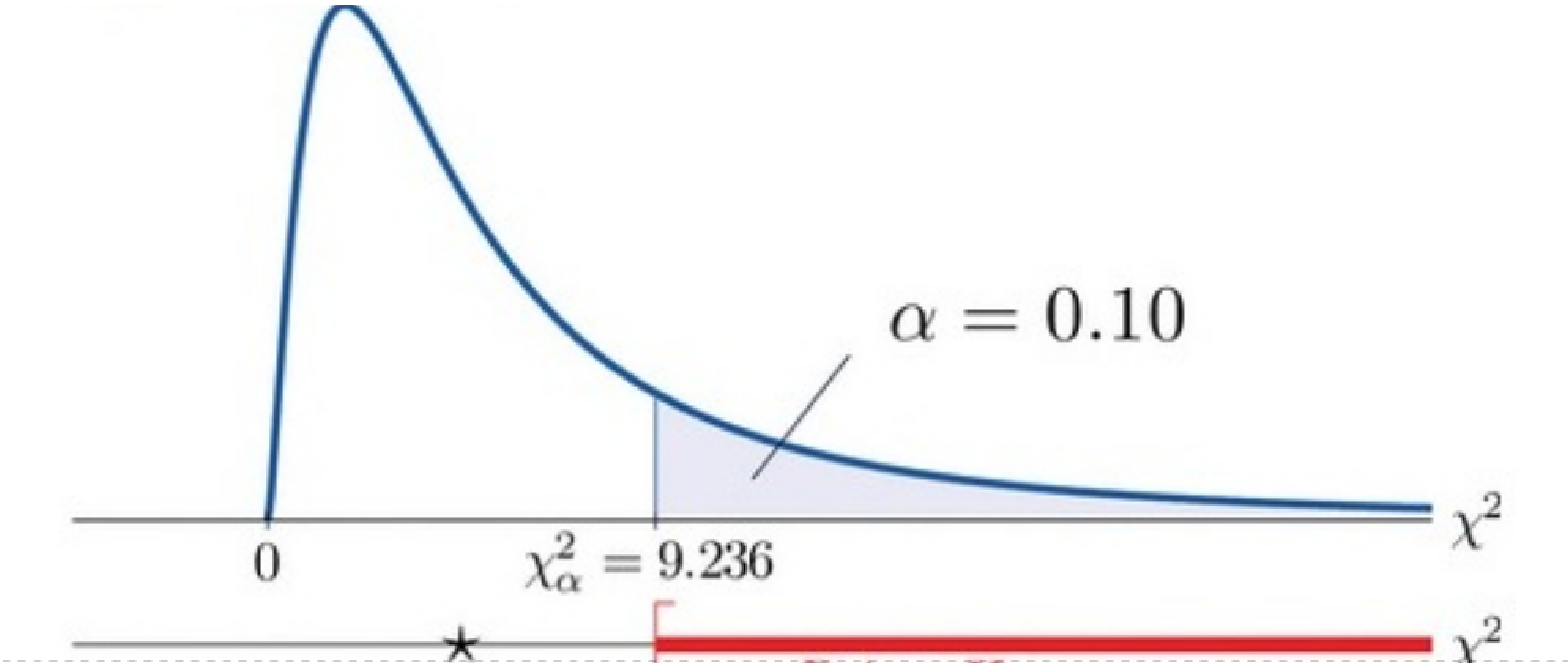
Die Value	Assumed Distribution	Observed Frequency
1	1/6	9
2	1/6	15
3	1/6	9
4	1/6	8
5	1/6	6
6	1/6	13

H0 : The die is fair
vs. Ha : The die is not fair

- Calculating the sum of square differences….

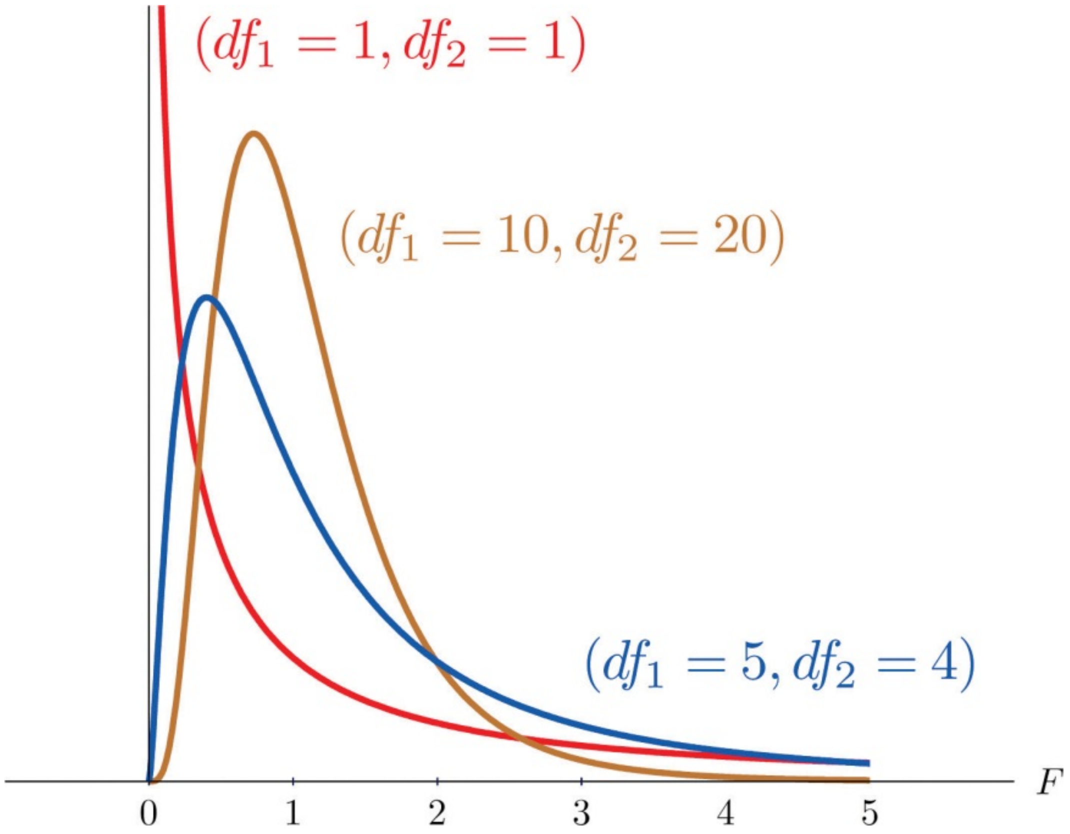
$$
\begin{aligned}
&\Sigma \frac{(O-E)^2}{E} \\
&= \frac{(-1)^2}{10} + \frac{5^2}{10} + \frac{(-1)^2}{10} + \frac{(-2)^2}{10} + \frac{(-4)^2}{10} + \frac{3^2}{10} \\
&= 0.1 + 2.5 + 0.1 + 0.4 + 1.6 + 0.9 \\
&= 5.6
\end{aligned}
$$

- Estimate critical value for 10% and df = 5



- Appears to be fair. We can generalize this to any discrete distribution

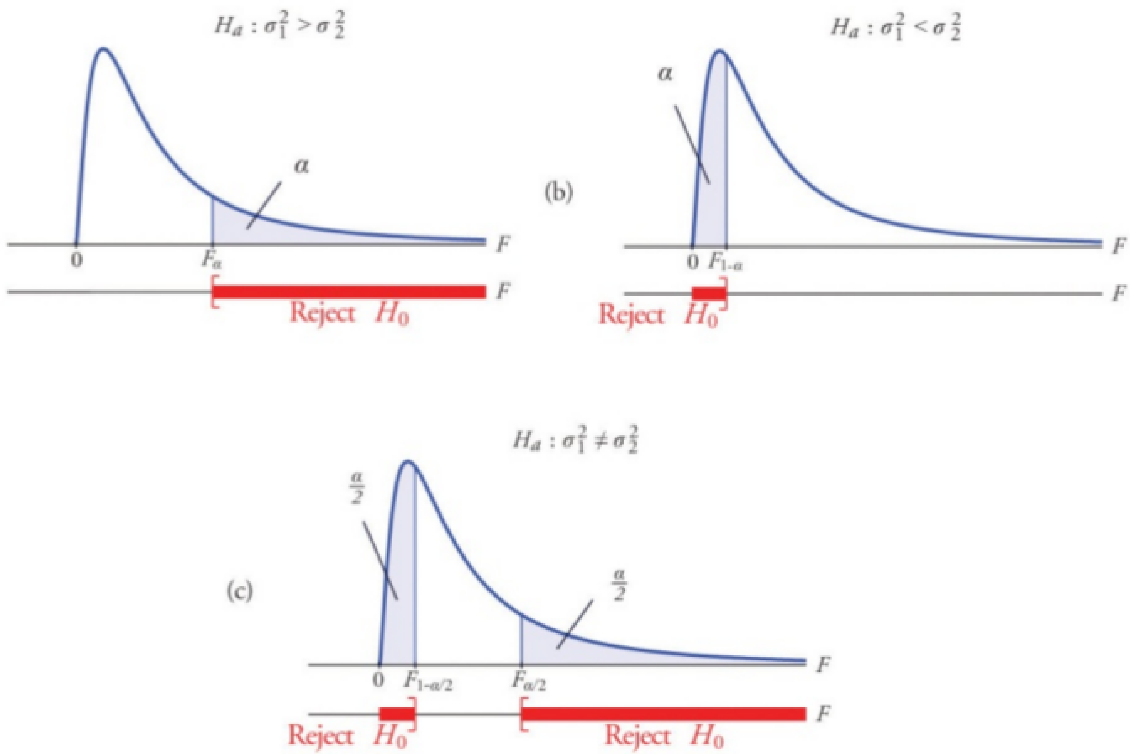
F-test for variances



- F-distribution is specified by a pair of degrees of freedoms (df1,df2)
- The parameter df1 is often referred to as the numerator degrees of freedom and the parameter df2 as the denominator degrees of freedom. They are not interchangeable.

Possible assumptions about Ha

- We tested the differences in sample means with the t-test, but now we are interested in a difference in the variance of two samples.
- H0: $\sigma_1^2 = \sigma_2^2$
- Right-tailed
- Left-tailed
- Two-tailed



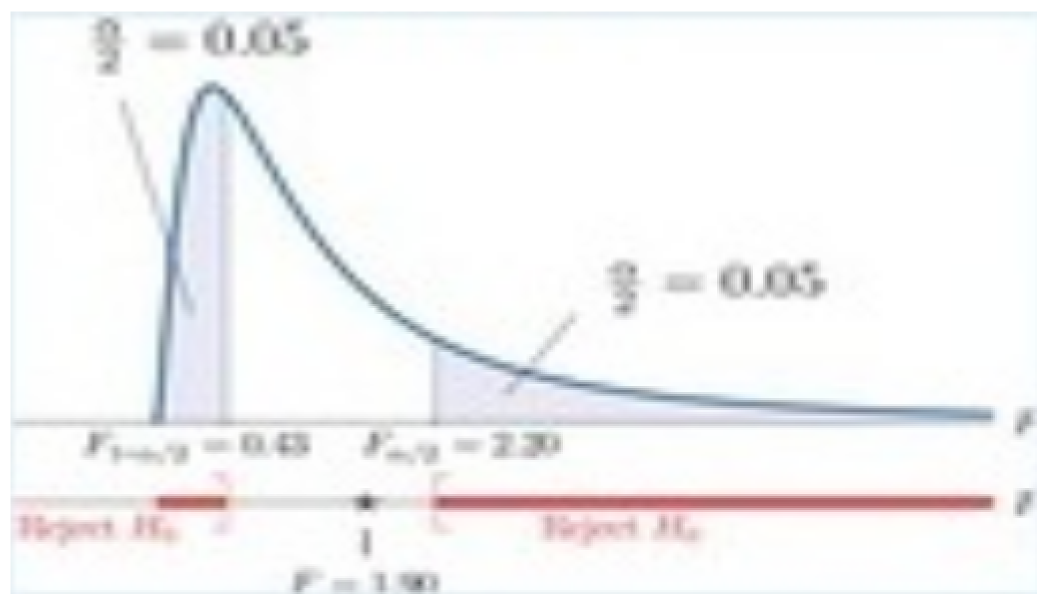
F-test example

- 2 types of Blood test stripes for glucose readings. Is their variance different at a 10% level of significance ?
- Type A: 16 samples, $s1^2 = 2.09$
- Type B: 21 samples, $s2^2 = 1.10$

- $H0 = \sigma_1^2 = \sigma_2^2$

  Vs $Ha = \sigma_1^2 \mathbin{/=} \sigma_2^2$ alpha= 0.1
- df1= 16-1   df2= 21-1

```
> qf(0.05,df1 = 15,df2 = 20,lower.tail = T)
[1] 0.4296391
> qf(0.05,df1 = 15,df2 = 20,lower.tail = F)
[1] 2.203274
```



- $F = \sigma 1^2 / \sigma 2^2$

  $= 2.09/1.1 = 1.9$

F-test in one-way ANOVA

- We compared two sample means, but now we are interested in comparing multiple sample means.

- K normal populations with possibly different means, μ_1, μ_2, ..., μ_K, but same variance σ^2

- We test:

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_K$$

vs. H_a : not all K population means are equal

- For the test K independent random samples are taken from the K normal populations.
- The K sample means, the K sample variances, and the K sample sizes estimated:

Population	Sample Size	Sample Mean	Sample Variance
1	n_1	$\bar{x}_1$	s_1^2
2	n_2	$\bar{x}_2$	s_2^2
$\vdots$	$\vdots$	$\vdots$	$\vdots$
K	n_K	$\bar{x}_K$	s_K^2

- Estimate quantities ->
- Notice that we use F-dist. to compare both variances

The **combined sample size:**

$$n = n_1 + n_2 + \cdots + n_K$$

The **mean of the combined sample** of all n observations:

$$\overline{x} = \frac{\Sigma x}{n} = \frac{n_1 \overline{x}_1 + n_2 \overline{x}_2 + \cdots + n_K \overline{x}_K}{n}$$

The **mean square for treatment**:

$$MST = \frac{n_1 \left(\overline{x}_1 - \overline{x}\right)^2 + n_2 \left(\overline{x}_2 - \overline{x}\right)^2 + \cdots + n_K \left(\overline{x}_K - \overline{x}\right)^2}{K-1}$$

The **mean square for error**:

$$MSE = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \cdots + (n_K - 1)s_K^2}{n - K}$$

MST can be thought of as the variance between the K individual independent random samples and **MSE** as the variance within the samples. This is the reason for the name "analysis of variance," universally abbreviated **ANOVA**. The adjective "one-way" has to do with the fact that the sampling scheme is the simplest possible, that of taking one random sample from each population under consideration. If the means of the K populations are all the same then the two quantities MST and MSE should be close to the same, so the null hypothesis will be rejected if the ratio of these two quantities is significantly greater than 1. This yields the following test statistic and methods and conditions for its use.

Test Statistic for Testing the Null Hypothesis that K Population Means Are Equal

$$F = \frac{MST}{MSE}$$

If the K populations are normally distributed with a common variance and if $H_0: \mu_1 = \cdots = \mu_K$ is true then under independent random sampling F approximately follows an F-distribution with degrees of freedom $df_1 = K-1$ and $df_2 = n - K$.

The test is right-tailed: H_0 is rejected at level of significance α if $F \geq F_\alpha$.

As always the test is performed using the usual five-step procedure.

Question

A random sample of major grade point averages (GPA) of 11 graduating seniors at a large university is selected for each of the four majors mathematics, English, education, and biology. Test, at the 5% level of significance, whether the data contain sufficient evidence to conclude that there are differences among the average major GPAs of these four majors.

- Sample Data:

Mathematics	English	Education	Biology
2.59	3.64	4.00	2.78
3.13	3.19	3.59	3.51
2.97	3.15	2.80	2.65
2.50	3.78	2.39	3.16
2.53	3.03	3.47	2.94
3.29	2.61	3.59	2.32
2.53	3.20	3.74	2.58
3.17	3.30	3.77	3.21
2.70	3.54	3.13	3.23
3.88	3.25	3.00	3.57
2.64	4.00	3.47	3.22

Major	Sample Size	Sample Mean	Sample Variance
Mathematics	$n_1 = 11$	$\bar{x}_1 = 2.90$	$s_1^2 = 0.188$
English	$n_2 = 11$	$\bar{x}_2 = 3.34$	$s_2^2 = 0.148$
Education	$n_3 = 11$	$\bar{x}_3 = 3.36$	$s_3^2 = 0.229$
Biology	$n_4 = 11$	$\bar{x}_4 = 3.02$	$s_4^2 = 0.157$

- State H_0 and Ha
- Estimate Critical values
- Df1 = k-1=4-1=3 , df2=n-k=44-4=40
- Estimate MST, and MSE, and test statistic
- F= MST/MSE
- Reject H0 or not

All steps neatly from the book

Solution:

- Step 1. The test of hypotheses is

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$
$$\text{vs. } H_a : \text{not all four population means are equal} \quad @ \alpha = 0.05$$

- Step 2. The test statistic is $F = MST / MSE$ with (since $n = 44$ and $K = 4$) degrees of freedom $df_1 = K - 1 = 4 - 1 = 3$ and $df_2 = n - K = 44 - 4 = 40$.

- Step 3. If we index the population of mathematics majors by 1, English majors by 2, education majors by 3, and biology majors by 4, then the sample sizes, sample means, and sample variances of the four samples in Table 11.17 "Difficulty Levels of College Majors" are summarized (after rounding for simplicity) by:

Major	Sample Size	Sample Mean	Sample Variance
Mathematics	$n_1 = 11$	$\bar{x}_1 = 2.90$	$s_1^2 = 0.188$
English	$n_2 = 11$	$\bar{x}_2 = 3.34$	$s_2^2 = 0.148$
Education	$n_3 = 11$	$\bar{x}_3 = 3.36$	$s_3^2 = 0.229$
Biology	$n_4 = 11$	$\bar{x}_4 = 3.02$	$s_4^2 = 0.157$

The average of all 44 observations is (after rounding for simplicity) $\bar{x} = 3.15$. We compute (rounding for simplicity)

$$MST = \frac{11(2.90 - 3.15)^2 + 11(3.34 - 3.15)^2 + 11(3.36 - 3.15)^2 + 11(3.02 - 3.15)^2}{4 - 1} = \frac{1.7556}{3} = 0.585$$

and

$$MSE = \frac{(11-1)(0.188) + (11-1)(0.148) + (11-1)(0.229) + (11-1)(0.157)}{44 - 4} = \frac{7.22}{40} = 0.181$$

so that

$$F = \frac{MST}{MSE} = \frac{0.585}{0.181} = 3.232$$

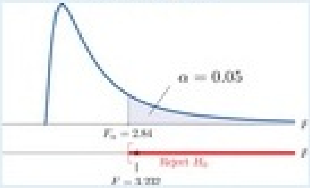
- Step 4. The test is right-tailed. The single critical value is (since $df_1 = 3$ and $df_2 = 40$) $F_\alpha = F_{0.05} = 2.84$. Thus the rejection region is $[2.84, \infty)$, as illustrated in Figure 11.12.

Figure 11.12

Note 11.36

"Example 8"

Rejection Region



- Step 5. Since $F = 3.232 > 2.84$, we reject H_0. The data provide sufficient evidence, at the 5% level of significance, to conclude that the averages of major GPAs for the four majors considered are not all equal.