CHAPTER 11: TWO POPULATION HYPOTHESIS TESTS

Tests of the Difference between two Normal Population Means: Dependent Samples

- Let's denote the means of two normally distributed populations by μ_X and μ_Y. Assume the random sample of $x_1, x_2, ..., x_n$ are drawn from the first population, and the random sample of $y_1, y_2, ..., y_n$ are drawn from the second population. Let's show the difference of observations by $d_i = x_i - y_i$, and the mean and standard deviation of these differences by $\bar{d}$ and s_d. Ifthe population distribution of these differences is assumed to be normal, then the following tests have a significance level α:

1. To test the null hypothesis

$$H_0 : \mu_x - \mu_y = D_0 \quad \text{or} \quad H_0 : \mu_x - \mu_y \leq D_0$$

against the alternative

$$H_1 : \mu_x - \mu_y > D_0$$

the decision rule is as follows

$$Reject \;\; H_0 \;\; if : \quad \frac{\bar{d} - D_0}{s_d / \sqrt{n}} > t_{n-1,\alpha}$$

2. To test the null hypothesis

$$H_0 : \mu_x - \mu_y = D_0 \quad \text{or} \quad H_0 : \mu_x - \mu_y \geq D_0$$

against the alternative

$$H_1 : \mu_x - \mu_y < D_0$$

the decision rule is as follows

$$Reject \;\; H_0 \;\; if : \quad \frac{\bar{d} - D_0}{s_d/\sqrt{n}} < -t_{n-1,\alpha}$$

3. To test the null hypothesis

$$H_0 : \mu_x - \mu_y = D_0$$

against the two-sides alternative

$$H_1 : \mu_x - \mu_y \neq D_0$$

the decision rule is as follows

$$Reject \;\; H_0 \;\; if: \quad \frac{\bar{d} - D_0}{s_d/\sqrt{n}} < -t_{n-1,\alpha/2}$$

$$\text{or} \qquad \frac{\bar{d} - D_0}{s_d/\sqrt{n}} > t_{n-1,\alpha/2}$$

where $t_{n-1,\alpha}$ is the number for which

$$P(t_{n-1} > t_{n-1,\alpha}) = \alpha$$

- Örnek: 10 ürün için 2 farklı reklam izleyicilere seyrettirilmiş, 24 saat sonra ise reklamların içeriği sorulmuştur. Reklamlar 'çok hatırlanan' ve 'az hatırlanan' diye kategorize edilmiş, bu iki grup içinde izleyicilerin reklam esnasındaki beyin aktivite miktarları bir indeks yardımıyla aşağıdaki gibi ölçülmüştür. Populasyonlardaki ortalama beyin aktivite miktarlarının birbirine eşit olduğu null hipotezini alternatif hipotez olan çok hatırlanan reklamlar için ortalama daha fazla beyin aktivitisi olduğu alternatif hipotezine karşı %5 ve %2.5 önem derecesinde test edelim

i	x_i	y_i	d_i	d_i^2
1	137	53	84	7,056
2	135	114	21	441
3	83	81	2	4
4	125	86	39	1,521
5	47	34	13	169
6	46	66	-20	400
7	114	89	25	625
8	157	113	44	1,936
9	57	88	-31	961
10	144	111	33	1,089

Örneklem ortalaması

$$\bar{d} = \frac{\sum_{i=1}^{n} d_i}{n} = \frac{210}{10} = 21$$

ve örneklem varyasyonu

$$s^2 = \frac{\sum_{i=1}^{n} d_i^2 - n\bar{d}^2}{n-1} = \frac{14,202^2 - 10 * 21^2}{9} = 1,088$$

ve standart sapması

$$s = \sqrt{1,0886} = 32.98$$

– Biz aşağıdaki hipotezi

$$H_0 : \mu_x - \mu_y = 0$$

şu alternetif hipoteze göre test etmek istiyoruz

$$H_1 : \mu_x - \mu_y > 0$$

– H_0'ı H_1'a karşı reddetmek için karar kuralında kullanılmak üzere

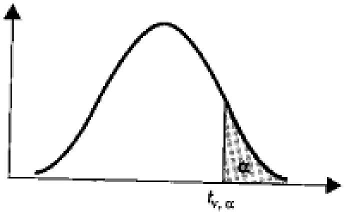
$$\frac{\bar{d} - D_0}{s/\sqrt{n}} = \frac{21 - 0}{32.98/\sqrt{10}} = 2.014$$

– Şimdi %5 ve %2.5 önem derecesinde testimizi yapalım. Student-t dağılımından bu önem derecelerine denk gelen n-1=10-1=9 serbestlik dereceli için değerler

$$t_{9,.05} = 1.833 \quad \text{and} \quad t_{9,.025} = 2.262$$

bizim bulduğumuz değer olan 2.014 değeri %5 önem derecesi değeri olan 1.833'den büyük olduğu için hipotezimizi bu önem derecesinde reddediyor, %2.5 önem derecesi değeri olan 2.262'den küçük olduğu için hipotezimizi %2.5 önem derecesinde reddedemiyoruz

TABLE 6 Cutoff points for the Student's t distribution



For selected probabilities, α, the table shows the values $t_{\nu,\alpha}$ such that $P(t_\nu > t_{\nu,\alpha}) = \alpha$, where t_ν is a Student's t random variable with ν degrees of freedom. For example, the probability is .10 that a Student's t random variable with 10 degrees of freedom exceeds 1.372.

ν	α				
	.100	.050	.025	.010	.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	[illegible]	[illegible]

Tests of the Difference Between Population Means: Independent Samples (Known Variances)

- Suppose n_x observation are drawn from a population with mean μ_X and variance σ^2_X, and n_y observation are drawn from a population with mean μ_Y and variance σ^2_Y, let the respective sample means are $\bar{x}$ and $\bar{y}$, then the following random variable has a standard normal distribution

$$Z = (\bar{x} - \bar{y}) = \frac{(\bar{x} - \bar{y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_Y^2}{n_y}}}$$

and the following tests have a significance level α:

1. To test the null hypothesis

$$H_0 : \mu_x - \mu_y = D_0 \quad \text{or} \quad H_0 : \mu_x - \mu_y \leq D_0$$

against the alternative

$$H_1 : \mu_x - \mu_y > D_0$$

the decision rule is as follows

$$Reject \;\; H_0 \;\; if: \quad \frac{\bar{x} - \bar{y} - D_0}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_Y^2}{n_y}}} > z_\alpha$$

2. To test the null hypothesis

$$H_0 : \mu_x - \mu_y = D_0 \quad \text{or} \quad H_0 : \mu_x - \mu_y \geq D_0$$

against the alternative

$$H_1 : \mu_x - \mu_y < D_0$$

the decision rule is as follows

$$Reject \;\; H_0 \;\; if: \quad \frac{\bar{x} - \bar{y} - D_0}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_Y^2}{n_y}}} < -z_\alpha$$

3. To test the null hypothesis

$$H_0 : \mu_x - \mu_y = 0$$

against the two-sided alternative

$$H_1 : \mu_x - \mu_y \neq D_0$$

the decision rule is as follows

$$Reject \;\; H_0 \;\; if: \quad \frac{\bar{x} - \bar{y} - D_0}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_Y^2}{n_y}}} < -z_{\alpha/2}$$

$$\text{or} \qquad \frac{\bar{x} - \bar{y} - D_0}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_Y^2}{n_y}}} > z_{\alpha/2}$$

• Örnek: İşyerlerinde kadın ve erkeklere verilen iş yükleri arasında fark olup olmadığının araştırılmasında, gözlemcilere fikirlerini 5 (çok fark var)'dan 1 (fark yok)'a kadar bir ölçekte belirtmeleri istenmmiştir.186 erkek gözlemcinin ortalaması 4.059, 172 kadın gözlemcinin ortalaması da 3.680 olmuştur. Önceki yıllardan populasyon standart sapmalarınında sırasıyla 0.839 ve 0.966 olduğu bilinmektedir. Örneklemlerin seçildiği kadın ve erkek gözlemci populasyonlarının değerlendirme ortalamalarının eşit olduğu null hipotezini, alternatif hipotez olan erkek gözlemcilerinn ortalamasının fazla olduğu hipotezine karşı test ediniz %1 önem derecesinde test ediniz.

– Biz aşağıdaki hipotezi (eğer erkekleri x, kadınları y ile gösterirsek)

$$H_0 : \mu_x - \mu_y = 0$$

şu alternetif hipoteze göre test etmek istiyoruz

$$H_1 : \mu_x - \mu_y > 0$$

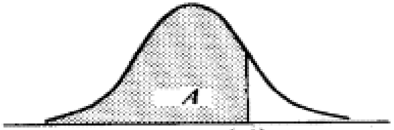
– H_0'ı H_1'a karşı reddetmek için kullanacağımız karar kuralı ise:
$$\frac{\bar{x}-\bar{y}}{\sqrt{\frac{\sigma_x^2}{n_x}+\frac{\sigma_Y^2}{n_y}}} > z_\alpha$$

- Verilenler: $\bar{x} = 4.059$ $\quad n_x = 186$ $\quad \sigma_x = .839$ $\quad \bar{y} = 3.680$ $\quad n_y = 172$ $\quad \sigma_y = .966$

$$\frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_Y^2}{n_y}}} > z_\alpha = \frac{4.059 - 3.680}{\sqrt{\frac{.839^2}{186} + \frac{.966^2}{172}}}$$

$= 3.95 > z_{0.01} = 2.33$ ise test reddedilir

Entry is area A under the standard normal curve from $-\infty$ to $z(A)$



$z(A)$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890

Tests of the Difference between two Population Means: Independent Samples, Unknown Population Variances Assumed to be Equal

- Remember the case where we don't know the population variances of the independent populations. If these unknown variances are assumed to be equal: $\sigma_x^2 = \sigma_y^2 = \sigma^2$, then we can use sample variances, $s_x^2 = s_y^2$, to estimate the common population variance σ^2. The estimate is

$$s_p^2 = \frac{(n_x - 1)s_x^2 + (n_y - 1)s_y^2}{n_x + n_y - 2}$$

since we use sample standart deviations, we refer to Student t-distributions

$$t = \frac{(\bar{x} - \bar{y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{s_p^2}{n_x} + \frac{s_p^2}{n_y}}}$$

and the following tests have a significance level α:

1. To test the null hypothesis

$$H_0 : \mu_x - \mu_y = 0 \quad \text{or} \quad H_0 : \mu_x - \mu_y \leq D_0$$

against the alternative

$$H_1 : \mu_x - \mu_y > D_0$$

the decision rule is as follows

$$Reject \;\; H_0 \;\; if : \quad \frac{\bar{x} - \bar{y} - D_0}{s\sqrt{\dfrac{n_x + n_y}{n_x n_y}}} > t_{n_x+n_y-2,\alpha}$$

2. To test the null hypothesis

$$H_0 : \mu_x - \mu_y = D_0 \quad \text{or} \quad H_0 : \mu_x - \mu_y \geq D_0$$

against the alternative

$$H_1 : \mu_x - \mu_y < D_0$$

the decision rule is as follows

$$Reject \;\; H_0 \;\; if : \quad \frac{\bar{x} - \bar{y} - D_0}{s\sqrt{\frac{n_x + n_y}{n_x n_y}}} < -t_{n_x+n_y-2,\alpha}$$

3. To test the null hypothesis

$$H_0 : \mu_x - \mu_y = D_0$$

against the two-sided alternative

$$H_1 : \mu_x - \mu_y \neq D_0$$

the decision rule is as follows

$$Reject \;\; H_0 \;\; if: \quad \frac{\bar{x}-\bar{y}-D_0}{s\sqrt{\frac{n_x+n_y}{n_x n_y}}} < -t_{n_x,n_y-2,\alpha}$$

$$\text{or} \quad \frac{\bar{x}-\bar{y}-D_0}{s\sqrt{\frac{n_x+n_y}{n_x n_y}}} > t_{n_x+n_y-2,\alpha}$$

$$P(t_{n_x,n_y-2,\alpha} > t_{n_x,n_y-2,\alpha}) = \alpha$$

- Örnek: Moderatörlerin, yönettikleri gruplarda ortaya çıkan fikirlerin sayısına etkisini inceleyen bir çalışmada, 4 kişilik moderatörlü ve moderatörsüz gruplar seçilmiştir. Moderatörlü gruplardaki ortaya çıkan ortalama fikir sayısı 78, (örneklem) standart sapması 24.4'tür. Moderatörsüz gruplardaki ortaya çıkan ortalama fikir sayısı 63.5, (örneklem) standart sapması 20.2.'dir. Populasyonların normal dağıldıklarını ve varyanslarının aynı olduğunu kabul edersek, populasyon ortalamalarının aynı olduğu null hipotezini moderatörlü grubun ortalamasının daha fazla olduğu alternatif hipoteze karşı %10 önem derecesinde test ediniz

– Biz aşağıdaki hipotezi(eğer moderatörlü grupları x, diğerlerini y ile gösterirsek)

$$H_0 : \mu_x - \mu_y = 0$$

şu alternetif hipoteze göre test etmek istiyoruz

$$H_1 : \mu_x - \mu_y > 0$$

– H_0'ı H_1'a karşı reddetmek için karar kuralı (decision rule)

$$Reject \;\; H_0 \;\; if: \quad \frac{\bar{x} - \bar{y} - D_0}{s\sqrt{\frac{n_x + n_y}{n_x n_y}}} > t_{n_x+n_y-2,\alpha}$$

- Verilenler: $\bar{x} = 78.0$ $n_x = 4$ $s_x = 24.4$ $\bar{y} = 63.5$ $n_y = 4$ $s_y = 20.2$
- İki grubun birbiriyle aynı olan populasyon standart sapmaları için bir tahminde bulunuyoruz

$$s^2 = \frac{(n_x - 1)s_x^2 + (n_y - 1)s_y^2}{n_x + n_y - 2} = \frac{(3)24.4^2 + (3)20.2^2}{4 + 4 - 2}$$
$$= 501.7$$

$$s = \sqrt{501.7} = 22.4$$

Şimdi ise tüm değerleri karar kuralında kullan-abilirirz

$$\frac{\bar{x}-\bar{y}}{s\sqrt{\frac{n_x+n_y}{n_x n_y}}} = \frac{78-63.5}{22.4\sqrt{\frac{8}{16}}} = 0.915$$

– Her biri 4 elamanlı iki örneklemimiz olduğu için, standart normal tablosundan %10'e denk gelen $student - t$ değeri için 4+4-2 değerine bakıyoruz

$$t_{n_x+n_y-2,\alpha} = t_{6,.10} = 1.44$$

dolayısıyla 0.915 bu sayıdan daha küçük olduğundan null hipotezini %10 önem düzeyinde alternatif tek yönlü hipoteze göre reddedemiyoruz

Tests of the Difference between Two Population Proportions (Large Samples)

- Remember the case where we were are interested in the difference between the proportion of successes in two populations. The distribution of the difference of the proportion of successes in two samples

$$Z = \frac{(\hat{p}_X - \hat{p}_Y) - (P_X - P_Y)}{\sqrt{\dfrac{\hat{p}_X(1-\hat{p}_X)}{n_x} + \dfrac{\hat{p}_Y(1-\hat{p}_Y)}{n_y}}}$$

If we want to test the hypothesis that population porportions are equal, we can assume that common value for these proportions, P_0

$$Z = \frac{(\hat{p}_X - \hat{p}_Y)}{\sqrt{\frac{P_0(1-P_0)}{n_x} + \frac{P_0(1-P_0)}{n_y}}} = \frac{(\hat{p}_X - \hat{p}_Y)}{\sqrt{P_0(1-P_0)\frac{n_x + \ldots}{n_x \ldots}}}$$

this common value further can be estimated by

$$\hat{P}_0 = \frac{n_x\hat{p}_X + n_y\hat{p}_Y}{n_x + n_y}$$

and the following tests have a significance level α:

1. To test the null hypothesis

$$H_0 : p_x - p_y = 0 \quad \text{or} \quad H_0 : p_x - p_y \leq 0$$

against the alternative

$$H_1 : p_x - p_y > 0$$

the decision rule is as follows

$$Reject \;\; H_0 \;\; if: \quad \frac{(\hat{p}_X - \hat{p}_Y)}{\sqrt{\hat{P}_0(1-\hat{P}_0)\frac{n_x+n_y}{n_x n_y}}} > z_\alpha$$

2. To test the null hypothesis

$$H_0 : p_x - p_y = 0 \quad \text{or} \quad H_0 : p_x - p_y \geq 0$$

against the alternative

$$H_1 : p_x - p_y < 0$$

the decision rule is as follows

$$Reject \;\; H_0 \;\; if : \quad \frac{(\hat{p}_X - \hat{p}_Y)}{\sqrt{\hat{P}_0(1 - \hat{P}_0)\frac{n_x + n_y}{n_x n_y}}} < -z_\alpha$$

3. To test the null hypothesis

$$H_0 : p_x - p_y = 0$$

against the two-sided alternative

$$H_1 : p_x - p_y \neq 0$$

the decision rule is as follows

$$Reject \;\; H_0 \;\; if: \quad \frac{(\hat{p}_X - \hat{p}_Y)}{\sqrt{\hat{P}_0(1-\hat{P}_0)\frac{n_x+n_y}{n_x n_y}}} < -z_{\alpha/2}$$

$$\text{or} \quad \frac{(\hat{p}_X - \hat{p}_Y)}{\sqrt{\hat{P}_0(1-\hat{P}_0)\frac{n_x+n_y}{n_x n_y}}} > z_{\alpha/2}$$

- Örnek: İngiltere'deki magazin reklamlarından 203 tanesinin 52 tanesi esprili bir şekilde yazılmıştır. Bağımsız başka bir örneklem olan 270 elemanlı Amerika'daki magazin reklamlalarından ise 56 tanesi esprili yazılmıştır. Bu iki reklam populasyonlarındaki esprili reklam oranının aynı olduğu hipotezi, alternatif iki yönlü hipoteze karşı test edildiğindeki p-değerini bulunuz

– Biz aşağıdaki hipotezi

$$H_0 : p_x - p_y = 0$$

şu alternetif hipoteze göre test etmek istiyoruz

$$H_1 : p_x - p_y \neq 0$$

– H_0'ı H_1'a karşı reddetmek için karar kuralı (decision rule)

$$\frac{(\hat{p}_X - \hat{p}_Y)}{\sqrt{\hat{P}_0(1 - \hat{P}_0)\frac{n_x + n_y}{n_x n_y}}} < -z_{\alpha/2} \quad \text{or} \quad > z_{\alpha/2}$$

– İki grubun populasyonları için ortak standart sapma bulmak için şu formulü kullanıyoruz

$$\hat{P}_0 = \frac{n_x \hat{p}_X + n_y \hat{p}_Y}{n_x + n_y}$$

– Soruda verilenler: $n_x = 203$ $\quad \hat{p}_X = 52/203 = .256$ $\quad n_y = 270$ $\quad \hat{p}_X = 56/270 = .207$, dolayısıyla

$$\hat{P}_0 = \frac{n_x \hat{p}_X + n_y \hat{p}_Y}{n_x + n_y} = \frac{203 * .256 + 270 * .207}{203 + 270} = .228$$

$$\frac{(\hat{p}_X - \hat{p}_Y)}{\sqrt{\hat{P}_0(1-\hat{P}_0)\frac{n_x + n_y}{n_x n_y}}} = \frac{.256 - .207}{\sqrt{.228 * .772 * \frac{203+270}{203*270}}} = 1.26$$

$*$ Standart normal tablosundan 1.26ya denk gelen $\alpha/2$ değeri ($z_{\alpha/2} = 1.26$) ise 0.8962. Dolayısıyla bu çift yönlü testin p-değeri .2*(1-0.8962)=2076'dır ve null hipotez olan populasyon ortalamalarının aynı olması hipotezi, %20.76'dan büyük herhangi bir önem derecesi ile reddedilebilir

Tests of Equality of the Variances between Two Normally Distributed Populations

- In this section we develop a procedure for testing the assumption that population variances from independent samples are equal. To perform such tests, we need to introduce F-distribution

The F-Distribution

- We have two independent random samples with n_x and n_y observations from two normal populations with variances σ_x^2 and σ_y^2. If the samle variances are s_x^2 and s_y^2, then the random variable

$$F = \frac{s_x^2/\sigma_x^2}{s_y^2/\sigma_y^2}$$

has an F distribution with numerator degrees of freedom $(n_y - 1)$ and denominator degrees of freedom $(n_y - 1)$

- This test is quite sensitive to normality
- We denote the number for which

$$P(F_{\upsilon_1,\upsilon_2} > F_{\upsilon_1,\upsilon_2,\alpha}) = \alpha$$

this is to say the probabilities for cut off points give the percentage of the data that in on the right hand side of the distribution

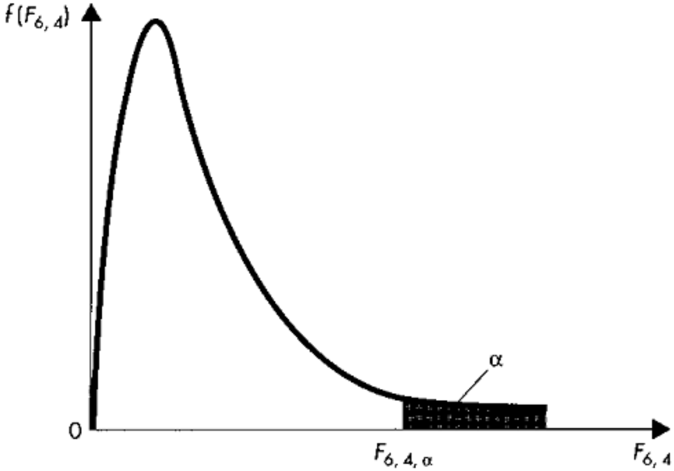
- Açıklayıcı Örnek: F dağılım tabloları her bir önem derecesi için ayrı ayrı hazırlanmıştır. Bunun sebebi ise F tablosuna bakmak istediğinizde,

$$F = \frac{s_x^2/\sigma_x^2}{s_y^2/\sigma_y^2}$$

bu formüldeki payın serbestlik derecesini tablonun kolonunda, paydanın serbestlik derecesini tablonun satırında bulmanız gerektiğidir. Dolayısıyla aynı tabloda bir de ayrıca önem derecesini görmek mümkün olamamaktadır. Yani birden fazla tabloya bakmamız gerekir

– F-dağılımının şekille göz atacak olursak

FIGURE 9.11 Probability density function of the F distribution with 6 numerator degrees of freedom and 4 denominator degrees of freedom; the probability is α that



– Örneğin paydaki örneklemin 11 elemanlı, paydadaki örneklemin 21 elemanlı olduğunu varsayarsak, bunların serbestlik dereceleri 10 ve 20 olur ki bunlar için %5 kesim noktası bu sayfayı takip eden 2. sayfadaki F-tablosundan aşağıdaki gibi bulunabilir

$$F_{10,20,.05} = 2.35$$

yani

$$P(F_{10,20} > 2.35) = .05$$

F - Distribution ($\alpha = 0.01$ in the Right Tail)

	Numerator Degrees of Freedom								
df_2 \ df_1	1	2	3	4	5	6	7	8	9
Denominator Degrees of Freedom									
1	4052.2	4999.5	5403.4	5624.6	5763.6	5859.0	5928.4	5981.1	6022.5
2	98.503	99.000	99.166	99.249	99.299	99.333	99.356	99.374	99.388
3	34.116	30.817	29.457	28.710	28.237	27.911	27.672	27.489	27.345
4	21.198	18.000	16.694	15.977	15.522	15.207	14.976	14.799	14.659
5	16.258	13.274	12.060	11.392	10.967	10.672	10.456	10.289	10.158
6	13.745	10.925	9.7795	9.1483	8.7459	8.4661	8.2600	8.1017	7.9761
7	12.246	9.5466	8.4513	7.8466	7.4604	7.1914	6.9928	6.8400	6.7188
8	11.259	8.6491	7.5910	7.0061	6.6318	6.3707	6.1776	6.0289	5.9106
9	10.561	8.0215	6.9919	6.4221	6.0569	5.8018	5.6129	5.4671	5.3511
10	10.044	7.5594	6.5523	5.9943	5.6363	5.3858	5.2001	5.0567	4.9424
11	9.6460	7.2057	6.2167	5.6683	5.3160	5.0692	4.8861	4.7445	4.6315
12	9.3302	6.9266	5.9525	5.4120	5.0643	4.8206	4.6395	4.4994	4.3875
13	9.0738	6.7010	5.7394	5.2053	4.8616	4.6204	4.4410	4.3021	4.1911
14	8.8616	6.5149	5.5639	5.0354	4.6950	4.4558	4.2779	4.1399	4.0297
15	8.6831	6.3589	5.4170	4.8932	4.5556	4.3183	4.1415	4.0045	3.8948
16	8.5310	6.2262	5.2922	4.7726	4.4374	4.2016	4.0259	3.8896	3.7804
17	8.3997	6.1121	5.1850	4.6690	4.3359	4.1015	3.9267	3.7910	3.6822
18	8.2854	6.0129	5.0919	4.5790	4.2479	4.0146	3.8406	3.7054	3.5971
19	8.1849	5.9259	5.0103	4.5003	4.1708	3.9386	3.7653	3.6305	3.5225
20	8.0960	5.8489	4.9382	4.4307	4.1027	3.8714	3.6987	3.5644	3.4567
21	8.0166	5.7804	4.8740	4.3688	4.0421	3.8117	3.6396	3.5056	3.3981
22	7.9454	5.7190	4.8166	4.3134	3.9880	3.7583	3.5867	3.4530	3.3458
23	7.8811	5.6637	4.7649	4.2636	3.9392	3.7102	3.5390	3.4057	3.2986
24	7.8229	5.6136	4.7181	4.2184	3.8951	3.6667	3.4959	3.3629	3.2560
25	7.7698	5.5680	4.6755	4.1774	3.8550	3.6272	3.4568	3.3239	3.2172
26	7.7213	5.5263	4.6366	4.1400	3.8183	3.5911	3.4210	3.2884	3.1818
27	7.6767	5.4881	4.6009	4.1056	3.7848	3.5580	3.3882	3.2558	3.1494
28	7.6356	5.4529	4.5681	4.0740	3.7539	3.5276	3.3581	3.2259	3.1195
29	7.5977	5.4204	4.5378	4.0449	3.7254	3.4995	3.3303	3.1982	3.0920
30	7.5625	5.3903	4.5097	4.0179	3.6990	3.4735	3.3045	3.1726	3.0665
40	7.3141	5.1785	4.3126	3.8283	3.5138	3.2910	3.1238	2.9930	2.8876
60	7.0771	4.9774	4.1259	3.6490	3.3389	3.1187	2.9530	2.8233	2.7185
120	6.8509	4.7865	3.9491	3.4795	3.1735	2.9559	2.7918	2.6629	2.5586
∞	6.6349	4.6052	3.7816	3.3192	3.0173	2.8020	2.6393	2.5113	2.4073

F - Distribution ($\alpha = 0.05$ in the Right Tail)

df_2 \ df_1 (Numerator Degrees of Freedom)	10	12	15	20	24	30	40	60	120	∞
1	241.88	243.91	245.95	248.01	249.05	250.10	251.14	252.20	253.25	254.31
2	19.396	19.413	19.429	19.446	19.454	19.462	19.471	19.479	19.487	19.496
3	8.7855	8.7446	8.7029	8.6602	8.6385	8.6166	8.5944	8.5720	8.5494	8.5264
4	5.9644	5.9117	5.8578	5.8025	5.7744	5.7459	5.7170	5.6877	5.6581	5.6281
5	4.7351	4.6777	4.6188	4.5581	4.5272	4.4957	4.4638	4.4314	4.3985	4.3650
6	4.0600	3.9999	3.9381	3.8742	3.8415	3.8082	3.7743	3.7398	3.7047	3.6689
7	3.6365	3.5747	3.5107	3.4445	3.4105	3.3758	3.3404	3.3043	3.2674	3.2298
8	3.3472	3.2839	3.2184	3.1503	3.1152	3.0794	3.0428	3.0053	2.9669	2.9276
9	3.1373	3.0729	3.0061	2.9365	2.9005	2.8637	2.8259	2.7872	2.7475	2.7067
10	2.9782	2.9130	2.8450	2.7740	2.7372	2.6996	2.6609	2.6211	2.5801	2.5379
11	2.8536	2.7876	2.7186	2.6464	2.6090	2.5705	2.5309	2.4901	2.4480	2.4045
12	2.7534	2.6866	2.6169	2.5436	2.5055	2.4663	2.4259	2.3842	2.3410	2.2962
13	2.6710	2.6037	2.5331	2.4589	2.4202	2.3803	2.3392	2.2966	2.2524	2.2064
14	2.6022	2.5342	2.4630	2.3879	2.3487	2.3082	2.2664	2.2229	2.1778	2.1307
15	2.5437	2.4753	2.4034	2.3275	2.2878	2.2468	2.2043	2.1601	2.1141	2.0658
16	2.4935	2.4247	2.3522	2.2756	2.2354	2.1938	2.1507	2.1058	2.0589	2.0096
17	2.4499	2.3807	2.3077	2.2304	2.1898	2.1477	2.1040	2.0584	2.0107	1.9604
18	2.4117	2.3421	2.2686	2.1906	2.1497	2.1071	2.0629	2.0166	1.9681	1.9168
19	2.3779	2.3080	2.2341	2.1555	2.1141	2.0712	2.0264	1.9795	1.9302	1.8780
20	2.3479	2.2776	2.2033	2.1242	2.0825	2.0391	1.9938	1.9464	1.8963	1.8432
21	2.3210	2.2504	2.1757	2.0960	2.0540	2.0102	1.9645	1.9165	1.8657	1.8117
22	2.2967	2.2258	2.1508	2.0707	2.0283	1.9842	1.9380	1.8894	1.8380	1.7831
23	2.2747	2.2036	2.1282	2.0476	2.0050	1.9605	1.9139	1.8648	1.8128	1.7570
24	2.2547	2.1834	2.1077	2.0267	1.9838	1.9390	1.8920	1.8424	1.7896	1.7330
25	2.2365	2.1649	2.0889	2.0075	1.9643	1.9192	1.8718	1.8217	1.7684	1.7110
26	2.2197	2.1479	2.0716	1.9898	1.9464	1.9010	1.8544	1.8027	1.7488	1.6906
27	2.2043	2.1323	2.0558	1.9736	1.9299	1.8842	1.8361	1.7851	1.7306	1.6717
28	2.1900	2.1179	2.0411	1.9586	1.9147	1.8687	1.8203	1.7689	1.7138	1.6541
29	2.1768	2.1045	2.0275	1.9446	1.9005	1.8543	1.8055	1.7537	1.6981	1.6376
30	2.1646	2.0921	2.0148	1.9317	1.8874	1.8409	1.7918	1.7396	1.6835	1.6223
40	2.0772	2.0035	1.9245	1.8389	1.7929	1.7444	1.6928	1.6373	1.5766	1.5089
60	1.9926	1.9174	1.8364	1.7480	1.7001	1.6491	1.5943	1.5343	1.4673	1.3893
120	1.9105	1.8337	1.7505	1.6587	1.6084	1.5543	1.4952	1.4290	1.3519	1.2539
∞	1.8307	1.7522	1.6664	1.5705	1.5173	1.4591	1.3940	1.3180	1.2214	1.0000

(df_2: Denominator Degrees of Freedom)

- Aşağıda bahsediyor olduğumuz testlerin mantğı, eğer bir örneklemn varyansı diğerinden yeterince fazla ise, populasyon varyanslarının birbirine eşit olduğu hipotezin reddedileceğidir

- We have two independent random samples with n_x and n_y observations from two normal populations with variances σ_x^2 and σ_y^2. If the sample variances are s_x^2 and s_y^2, and if s_x^2 is bigger than

s_y^2, then remember that the random variable

$$F = \frac{s_x^2/\sigma_x^2}{s_y^2/\sigma_y^2}$$

has an F distribution with numerator degrees of freedom $(n_y - 1)$ and denominator degrees of freedom $(n_y - 1)$. Then the following tests have significance level α

1. To test the null hypothesis

$$H_0 : \sigma_x^2 = \sigma_Y^2 \quad \text{or} \quad H_0 : \sigma_x^2 \leq \sigma_Y^2$$

against the alternative

$$H_1 : \sigma_x^2 > \sigma_Y^2$$

the decision rule is as follows

$$Reject \;\; H_0 \;\; if : \quad \frac{s_x^2}{s_y^2} > F_{n_x-1,n_y-1,\alpha}$$

2. To test the null hypothesis

$$H_0 : \sigma_x^2 = \sigma_Y^2$$

against the alternative

$$H_1 : \sigma_x^2 \neq \sigma_Y^2$$

the decision rule is as follows

$$Reject \;\; H_0 \;\; if : \quad \frac{s_x^2}{s_y^2} > F_{n_x-1, n_y-1, \alpha/2}$$

- Örnek: AAA kredi notuna sahip 17 hisse senedinin getiri varyansı 123.35, CCC kredi notuna sahip 11 hisse senedinin getiri varyansı 8.02 ise, aşağıdaki null hipotezini

$$H_0 : \sigma_x^2 = \sigma_Y^2$$

şu alternatif hipoteze göre %2 önem derecesinde test ediniz

$$H_1 : \sigma_x^2 \neq \sigma_Y^2$$

– H_0'ı H_1'a karşı reddetmek için karar kuralı (decision rule)

$$\frac{s_x^2}{s_y^2} > F_{n_x-1,n_y-1,\alpha/2}$$

– Soruda verilenler: $n_x = 17 \quad s_x^2 = 123.35 \quad n_y = 11 \quad s_y^2 = 8.02$, dolayısıyla

$$\frac{s_x^2}{s_y^2} = \frac{123.35}{8.02} = 15.38$$

– İki yönlü bir hipotezi %2 önem derecesinde test etmek istersek, dağılımın her bir tarafından %1 atacağımızdan; ayrıca payın serbestlik derecesi 16, paydanınki 10 olduğundan

$$F_{16,10,.01} = 4.53$$

dolayısıyla 15.38 değeri bu değerden büyük olduğundan null hipotezini %2 önem düzeyinde reddediyoruz