İlişkisiz İki Örneklem Testleri: Fisher Tam Olasılık ve Mann Whiney U Testleri

Nihat Tak

2023-05-11

Fisher Tam Olasılık Testi

- χ^2 testleri gözlerdeki frekansların küçük olması halinde,
- Gözlerden birinde veya daha fazlasında 5 ve daha az beklenen frekans varsa,
- 2*2 lik tablolarda birim sayısı 20 dan küçük ise(n<20)
 - İnceleyeceğimiz χ^2 hesapları bizi yanlış sonuçlara ulaştırır.
 - Bu durumda bilinen χ^2 yerine Fisher Tam Olasılık Testi kullanılması öngörülür.

Varsayımları

- Örneklemler rastsal ve birbirinden bağımsızdır.
- Her gözlem yalnızca bir kategoride yer almalıdır.
- Ancak bu test n in 20 veya daha küçük olduğu hem A+B hem de C+A nın 15 den büyük olmaması durumunda kullanılabilmektedir.

Yöntem:

- Fisher' in testi, gözlenen 2x2 lik bir tablonun kesin önem seviyesinin hesaplanmasına dayanır.
- Bu teste göre sıra ve sütun marjinal (kenar) toplamları ve (n) değişmemek koşulu ile tablo içi hücre değerlerinin değişik kombinasyonları için söz konusu olasılıkları hesaplayarak toplanır. Bu son elde edilen değer gözlenen ve daha ekstrem tabloların elde edilme olasılığıdır. Toplamı (n) olan hücreleri a,b,c,d olan bir tablonun elde edilme olasılığı:
- Eğer sıfır hipotezi doğru ise en fazla elde edilen tablo kadar a değeri elde etme olasılığı esas alınır.

Hipotez Testi Aşamaları

1. Adım Hipotezler kurulur.

 p_1 =1. kitleden ilgilenen özelliklerin oranı

 p_2 =2. kitleden ilgilenilen özelliğin oranı

$$H_0: p_1 = p_2, H_0: p_1 \le p_2, H_0: p_1 \ge p_2$$

$$H_a: p_1 \neq p_2, H_a: p_1 > p_2, H_a: p_1 < p_2$$

2. Adım Test istatistiği hesaplanır.

Gruplar arası fark var mı?

$$p = \frac{\binom{A+B}{A}\binom{C+D}{C}}{\binom{N}{A+C}}$$

	GRUPLAR	OLGI	TOPLAM	
		(-)	(+)	
P	1	(p ₁) A	В	A+B
Ļ	2	(p ₂) C	D	C+D
	TOPLAM	A+C	B+D	N

Değişkenler arası fark var mı?

$$p = \frac{\binom{A+C}{A}\binom{B+D}{B}}{\binom{N}{A+B}}$$

GRUPLAR	OLGULAR		TOPLAM
	(-)	(+)	
1	(p ₁)A	(p ₂)B	A+B
2	С	D	C+D
TOPLAM	A+C	B+D	N
	<u> </u>	<u> </u>	6

3. Adım Kritik değer bulunur.

Bu test için kritik değer soruda verilen hata payı olan α dır.

4. Adım Karar verilir.

Çift yönlü test: $p < \alpha/2$ ise H_0 reddedilir.

Tek yönlü test: $p < \alpha$ ise H_0 reddedilir.

Örnek 20 evli bayandan 13'ünün 2 veya daha az çocuğu vardır. Bunlara bir işte çalışıpçalışmadıkları soruldu, 10'u çalışıyordu.Geriye kalan 7 bayan üç veya daha fazla çocuklu idiler, bunların ise yalnızca 1'i çalışıyordu. Bu verilere bakarak 2 veya daha az çocuğa sahip olanlarda, çalışan evli bayanların oranı çalışmayan evli bayanlardan daha mı büyüktür?

GRUPLAR	OLGULAR		TOPLAM		
	2 veya az	3 vaya daha fazla			
Çalışan	(p ₁)10	1	11		
Çalışmayan	(p ₂)3	6	9		
TOPLAM	13	7	20		

1. Adım Hipotezler kurulur.

 H_0 : $p_1 \le p_2$ (2 veya daha az çocuğa sahip olanlarda, çalışan evli bayanların oranı ile çalışmayan evli bayanların oranı arası fark yoktur.)

 H_a : $p_1 > p_2$ (2 veya daha az çocuğa sahip olanlarda, çalışan evli bayanların oranı çalışmayan evli bayanlardan daha büyüktür)

2.Adım Test istatistiği hesaplanır.

GRUPLAR	OLG	TOPLAM		
	2 veya az	3 veya daha fazla		
Çalışan	10	1	11	
Çalışmay an	3	6	9	
TOPLAM	13	7	20	

GRUPLAR	OLGULAR		TOPLAM
	2 veya az	3 veya daha fazla	
Çalışan	11	0	11
Çalışmay an	2	7	9
TOPLAM	13	7	20

I. tablonun oluşma olasılığı

$$p = \frac{\binom{A+B}{A}\binom{C+D}{C}}{\binom{N}{A+C}}$$
$$p_a = \frac{\binom{11}{10}\binom{9}{3}}{\binom{20}{10}} = 0,0476$$

$$p_b = \frac{\binom{11}{11}\binom{9}{2}}{\binom{20}{13}} = 0,000464$$

$$p = p_a + p_b = 0.0119195 + 0.000464 = 0.0123839$$

3. Adım Kritik değer $\alpha = 0.05$ olarak kullanılır.

4.Adım Karar verilir.

p değeri α =0,05 değerinden daha küçük olmadığı için H_0 reddedilemez.

Büyük Örneklem Yaklaşımı

Yeterince büyük örneklemler için normal yaklaşımı kullanabiliriz. Aynı zamanda bu büyük örneklem için sıfır hipotezini test ederken ki kare testi de uygulanabilinir.

$$z = \frac{[A/(A+B)] - [C/(C+D)]}{\sqrt{\hat{p}(1-\hat{p})\frac{1}{(A+B)} + \frac{1}{(C+D)}}}$$
$$\hat{p} = \frac{(A+B)}{N}$$

Örnek Bir araştırmacı 20. gebelik haftasının altında gebeliği olan 224 gebe kadın üzerinde araştırma yapmıştır. Tüm gebelerden kan örnekleri alınıp incelenmiştir. Aşağıdaki sonuçlar elde edilmiştir.

YAŞ	OLGI	TOPLAM	
ARALIĞI	Toksoplazma	Toksoplazma	
	+	-	
18 - 30	58	131	189
31 - +	19	36	55
TOPLAM	77	187	244

Bu verilere göre 18-30 yaş arası gebeler ile 31 yaş veya daha üstü gebeler toksoplazma virüsü bulundurmaları bakımından farklı mıdır?

1. Adım Hipotezler kurulur.

 H_0 : $p_1 = p_2$ (Bu verilere göre 18-30 yaş arası gebeler ile 31 yaş veya daha üstü gebeler toksoplazma virüsü bulundurmaları bakımından fark yoktur.)

 H_a : $p_1 \neq p_2$ (Bu verilere göre 18-30 yaş arası gebeler ile 31 yaş veya daha üstü gebeler toksoplazma virüsü bulundurmaları bakımından farklıdır.)

2. Adım Test istatistiği hesaplanır.

YAŞ ARALIĞI	OLG	TOPLAM		
	Toksoplazma Toksoplazma			
	+			
18 - 30	58(A)	131(B)	189	
31 - +	19(C)	36(D)	55	
TOPLAM	77	187	244	

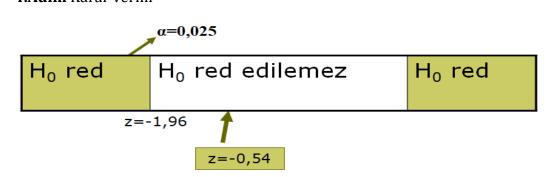
$$z = \frac{\left[A/(A+B)\right] - \left[C/(C+D)\right]}{\sqrt{\hat{p}(1-\hat{p})\left[\frac{1}{(A+B)} + \frac{1}{(C+D)}\right]}} = \frac{(58/189) - (19/55)}{\sqrt{0,3155 * 0,6845(0,00529 + 0,01818)}} = \frac{-0,0386}{0,071189} = -0,5422$$

$$\hat{p} = \frac{(A+C)}{N} \qquad \qquad \hat{p} = \frac{77}{244} = 0,3155$$

3.Adım Kritik tablo değeri bulunur.

$$Z_{\alpha/2} = Z_{0.025} = 1.96$$

4.Adım Karar verilir



P(z=-0.54) < 0.2646 olduğundan bu değer $\alpha/2=0.025$ ten küçük olmadığı için H_0 hipotezi reddedilemez

Bu verilere göre 18-30 yaş arası gebeler ile 31 yaş veya daha üstü gebeler toksoplazma virüsü bulundurmaları bakımından fark yoktur.

Örnek Bir araştırmacı 20. gebelik haftasının altında gebeliği olan 224 gebe kadın üzerinde araştırma yapmıştır. Tüm gebelerden kan örnekleri alınıp incelenmiştir. Aşağıdaki sonuçlar elde edilmiştir.

YAŞ	OLG	TOPLAM	
ARALIĞI	Toksoplazma	Toksoplazma	
	+	-	
18 - 30	58	131	189
31 - +	19	36	55
TOPLAM	77	187	244

Bu verilere göre 18-30 yaş arası gebeler arasında toksoplazma virüsü bulundurmaları ve bulundurmamaları bakımından farklı mıdır?

1.AdımHipotezler kurulur

 H_0 : $p_1 \le p_2$ (Bu verilere göre 18-30 yaş arası gebeler arasında toksoplazma virüsü bulundurmaları ve bulundurmamaları bakımından fark yoktur.)

 H_a : $p_1>p_2$ (Bu verilere göre 18-30 yaş arası gebeler arasında toksoplazma virüsü bulundurmaları ve bulundurmamaları bakımından farklıdır.)

2. Adım Test istatistiği hesaplanır

YAŞ ARALIĞI	OLG	TOPLAM	
	Toksoplazma + -		
18 - 30	58(A)	131(B)	189
31 - +	19(C)	36(D)	55
TOPLAM	77	187	244

$$z = \frac{\left[A/(A+C)\right] - \left[B/(B+D)\right]}{\sqrt{\hat{p}(1-\hat{p})\left[\frac{1}{(A+C)} + \frac{1}{(B+D)}\right]}} = -1.048$$

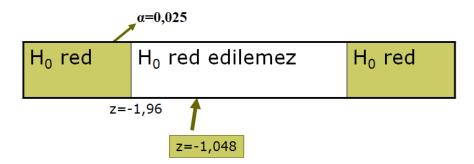
$$\hat{p} = \frac{(A+B)}{N}$$

$$\hat{p} = \frac{77}{244} = 0.774$$

3. Adım Kritik tablo değeri bulunur

$$Z_{\alpha/2} = Z_{0.025} = 1.96$$

4. Adım Karar verilir



P(z=-0.54)<0.4306 olduğundan bu değer $\alpha/2=0.025$ ten küçük olmadığı için H_0 hipotezi red edilemez

Bu verilere göre 18-30 yaş arası gebeler arasında toksoplazma virüsü bulundurmaları ve bulundurmamaları bakımından fark yoktur.

Özet

- -Fisher Tam Olasılık Testi, iki değişken arası bağımsızlığın ve homojenliğin belirlenmesin de kullanılır.
- -Gözlerden birinde veya daha fazlasında, 5 ve daha az beklenen frekans varsa, 2*2 lik tablolarda birim sayısı 20 den küçük ise(n<20) Ki Kare testi yerine Fisher Tam Olasılık Testi kullanılır.

Mann-Whitney U Testi

- İki bağımsız örnekleme ilişkin parametrik olmayan tekniklerden birisidir.
- Bu test Mann-Whitney Wilcoxon Testi olarak da adlandırılmaktadır.
- Mann-Whitney U testi, test istatistiği olarak sıra sayıları toplamlarını kullanarak, iki örneklem dağılımının eşit olup olmadığına karar veren bir uyum iyiliği testtir.

Varsayımlar

- Veriler x1, x2,...,xn1 ve y1, y2,...,yn2 gibi iki örneklemden oluşur.
- Örneklemler bağımsızdır.
- Gözlenen değişkenler sürekli rasgele değişkenlerdir.
- Ölçme düzeyi en az sıralayıcıdır.

Adımları

1.Adım Hipotezler kurulur

(iki yanlı test)

 H_0 : Kitleler benzer dağılımlara sahiptir.

 H_1 : Kitlelerin dağılımları birbirinden farklıdırlar.

(tek yanlı test)

 H_0 : Kitleler benzer dağılımlara sahiptir.

 H_1 : x'ler y'lerden daha küçük olma eğilimindedir.

(tek yanlı test)

 H_0 : Kitleler benzer dağılımlara sahiptir.

 H_1 : x'ler y'lerden daha büyük olma eğilimindedir.

2. Adım Test istatistiği hesaplanır

Adım1: İki örneklemdeki verilere, hangi örnekleme ait olduklarına bakılmaksızın, büyüklük sırasına göre sıra sayıları verilir.

Adım2: Birinci örnekleme ait sıra sayıları toplamı R_1 , ikinci örnekleme ait sıra sayıları toplamı R_2 , ve örneklem birim sayıları n_1 ve n_2 olmak üzere, u_1 ve u_2 istatistikleri hesaplanır:

$$u_1 = n_1 n_2 + (n_1(n_1 + 1)/2) - R_1$$

$$u_2 = n_1 n_2 + (n_2(n_2 + 1)/2) - R_2$$

$$u_1 + u_2 = n_1 n_2$$

$$n = n_1 + n_2$$

$$R_1 + R_2 = n(n + 1)/2$$

 u_1 ve u_2 istatistiklerinden küçük olan değer test istatistiği olarak kullanılır

3. Adım Kritik değer bulunur

Mann-Whitney U Testi tablosundan W_{α,n_1,n_2} değeri bulunur.

4. Adım Karar verilir

Eğer $min(u_1, u_2) < W_{n_1,n_2}$ ise H_0 reddedilir.

Büyük Örneklem

- Örneklem büyüklüğü n_1 ve/veya $n_2 > 20$ olması durumunda, W_{α,n_1,n_2} kritik değer tablodan belirlenemez.
- Büyük örneklem/örneklemler durumunda, z dönüşümü kullanılır.

$$\left. \begin{array}{l} \mathsf{E}(\mathsf{u}) \! = \! \mathsf{n_1} \mathsf{n_2} / 2 \\ \mathsf{var}(\mathsf{u}) \! = \! \mathsf{n_1} \mathsf{n_2} (\mathsf{n_1} \! + \! \mathsf{n_2} \! + \! 1) / 12 \end{array} \right\} \ z = \frac{\min(u_1, u_2) - n_1 n_2 / 2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1) / 12}}$$

 Aynı değerli gözlemler söz konusu olduğunda gruplar içerisinde aynı değerli gözlemler test istatistiğini etkilemez. Ancak gruplar arasındaki aynı değerli gözlemler, test istatistiğini etkiler. Bu durumda varyans için düzeltme terimi kullanılır. t_i :tekrarlanan gözlem değerlerinin sayıları

$$D.T. = \frac{n_1 n_2 (\Sigma t_i^2 - \Sigma t_i)}{12(n_1 + n_2)(n_1 + n_2 - 1)}$$

$$z = \frac{\min(u_1, u_2) - n_1 n_2 / 2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1) / 12 - D.T.}}$$

Örnek Trafik suçlarına verilen cezaların arttırılmasıyla birlikte Fevzipaşa Bulvarı üzerinde hatalarından dolayı cezaya çarptırılan 6 özel oto sürücüsü ile 8 taksi şoförünün ulaşmış oldukları toplam ceza puanları verilmektedir. Bu konuda taksi soförleri ile özel oto sürücüleri arasında fark var mıdır?(α =0.05)

Öz.oto sür.:20 35 25 10 10 30

Taks.sür. :40 15 60 50 35 50 10 45

1.Adım Hipotezler kurulur

 H_0 : Özel oto sürücülerinin aldıkları ceza puanları ile taksi sürücülerinin aldıkları ceza puanları birbirinden farklı değildir. (Kitleler benzer dağılımlara sahiptir.)

 H_1 : Özel oto sürücülerinin aldıkları ceza puanları ile taksi sürücülerinin aldıkları ceza puanları birbirinden farklıdır. (Kitlelerin dağılımları birbirinden farklıdırlar.)

2. Adım Test istatistiği hesaplanır

özel	Sıra no	Ort. Sıra no	taksi	Sıra no	Ort
					Sıra no
20	5	5	40	10	10
35	8	8.5	15	4	4
25	6	6	60	14	14
10	1	2	50	12	12.5
10	2	2	35	9	8.5
30	7	7	50	13	12.5
	$R_1 = 30,5$	R ₂ =74,5	10	3	2
			45	11	11

$$R_1=5+8,5+...+7=30,5$$
 $n=n_1+n_2$
 $R_2=10+4+...+11=74,5$ $R_1+R_2=n(n+1)/2$
 $u_1=6.8+6.7/2-30,5=38,5$
 $u_2=6.8+8.9/2-74,5=9,5$
 $(u_1+u_2)=n_1n_2$
 $min(u_1,u_2)=9,5$

3. Adım Kritik değer bulunur

$$w_{0,025;6,8} = 9$$

4. Adım Karar verilir

 $min(u1, u2) = 9.5 > 9 = w_{0.025:6.8}$ olduğundan H_0 'ı reddecek yeterli kanıt yoktur denir.

Özel oto sürücülerinin aldıkları ceza puanları ile taksi sürücülerinin aldıkları ceza puanları birbirinden farklı değildir.

ÖrnekRasgele seçilen 20 öğrenci rasgele 2 gruba ayrılmıştır. Yabancı dil dersinde, 1. gruptaki 12 öğrenci A tekniğiyle, 2. gruptaki 8 öğrenci de B tekniğiyle eğitilmişlerdir. Bu öğrencilerin genel bir testte aldıkları tabloda verilmektedir. Uzmanlar A tekniğiyle eğitilen öğrencilerin başarı puanlarının B tekniğiyle eğitilen öğrencilerin başarı puanlarından genel olarak daha düşük olduğunu iddia etmektedir.

1.Adım Hipotezler kurulur

 H_0 : A tekniğiyle eğitilen öğrencilerin başarı puanları ile B tekniğiyle eğitilen öğrencilerin başarı puanları birbirinden farklı değildir. (Kitleler benzer dağılımlara sahiptir.)

 H_1 : A tekniğiyle eğitilen öğrencilerin başarı puanları B tekniğiyle eğitilen öğrencilerin başarı puanlarından genel olarak daha düşüktür. (x'ler y'lerden daha küçük olma eğilimindedir.)

2. Adım Test istatistiği hesaplanır

A tekn.	Sıra no	B tekn.	Sıra no	
30	4	37	5	
20	3	80	16	
40	7	82	17	
60	14	59	13	
41	8	38	6	
50	10	91	20	
90	19	88	18	
15	2	68	15	
47	9			
56	12			
12	1			
51	11			

$$\begin{array}{lll} R_1 = 4 + 3 + \ldots + 11 = 100 & n = n_1 + n_2 \\ R_2 = 5 + 16 + \ldots + 15 = 110 & R_1 + R_2 = n(n+1)/2 \\ u_1 = 8.12 + 12.13/2 - 100 = 74 \\ u_2 = 8.12 + 8.9/2 - 110 = 22 \\ & (u_1 + u_2) = n_1 n_2 \\ \\ \\ min(u_1, u_2) = 22 \end{array}$$

3. Adım Kritik değer bulunur

$$W_{0.05:12.8} = 27$$

4. Adım Karar verilir

 $min(u_1, u_2) = 22 < 27 = w_{0,05;12,8}$ olduğundan H_0 red edilir.

A tekniğiyle eğitilen öğrencilerin başarı puanları B tekniğiyle eğitilen öğrencilerin başarı puanlarından genel olarak daha düşüktür.

Örnek Eskişehir' de aynı gün kurulan Yeni pazar ve Vişnelik pazarında benzer kalitedeki domateslerin fiyatları çeşitli sergilerden derlenerek elde edilmiştir ve tabloda verilmektedir. Domates fiyatlarının genel olarak Vişnelik pazarında daha ucuz olduğunu söylemek mümkün müdür?

Yeni Pazar (Kuruş)	Ort.Sır a no			Vişnelik Pazarı (Kuruş)	Ort.Sıra no		
600	25	600	25	550	17	350	2
650	32	650	32	500	12.5	350	2
700	38.5	550	17	350	2	400	5.5
600	25	550	17	450	8.5	450	8.5
575	20.5	600	25	475	10	500	12.5
600	25	650	32	400	5.5	500	12.5
667	36	700	38.5	550	17		
550	17	667	36	500	12.5		
625	29	650	32	600	25		
750	40.5	667	36	400	5.5		
600	25	650	32	575	20.5		
750	40,5			400	5.5		

1.Adım Hipotezler kurulur

 H_0 : İki pazarda da domates fiyatları birbirinden farklı değildir. (Kitleler benzer dağılımlara sahiptir.)

 H_1 : Vişnelik pazarında domates daha ucuzdur. (x'ler y'lerden daha küçük olma eğilimindedir.)

2. Adım Test istatistiği hesaplanır

$$\begin{array}{lll} R_1 = 25 + 32 + \ldots + 32 = 676, 5 & n = n_1 + n_2 \\ R_2 = 17 + 12.5 + \ldots + 12.5 = 184, 5 & R_1 + R_2 = n(n+1)/2 \\ u_1 = 23.18 + 23.24/2 - 676, 5 = 400, 5 \\ u_2 = 23.18 + 18.19/2 - 184, 5 = 13, 5 \\ & (u_1 + u_2) = n_1 n_2 \end{array}$$

 $min(u_1,u_2)=13,5$ $n_1>20$ olduğundan z dönüşümü yapılır.

z dönüşümü için

$$z = \frac{\min(u_1, u_2) - n_1 n_2 / 2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1) / 12 - D.T.}} \quad D.T. = \frac{n_1 n_2 (\sum t_i^2 - \sum t_i)}{12(n_1 + n_2)(n_1 + n_2 - 1)}$$

hesaplanır. Öncelikle

$$t_1=7, t_2=5, ..., t_{11}=4$$

$$z = \frac{\min(u_1, u_2) - n_1 n_2 / 2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1) / 12 - D.T.}} = \frac{13.5 - 23.18 / 2}{\sqrt{23.18(23 + 18 + 1) / 12 - 16}} = -5.11$$

3. Adım Kritik değer bulunur

4. Adım Karar verilir

$$|z|=5.11>1.645=z_{0.05}$$
 olduğundan H_0 reddedilir.

Yani %5 önem düzeyinde vişnelik pazarında domates daha ucuzdur.

Çİft Yönlü Test İçin $\alpha=0.05$, Tek Yönlü Test İçin $\alpha=0.025$

n_1	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	9	12	14	16	19	21	23	26	28	31	33	35	38	40	42	45	47	49
4	12	15	18	21	24	27	30	33	36	39	42	45	48	50	53	56	59	62
5	14	18	21	25	29	32	36	39	43	47	50	54	57	61	65	68	72	75
6	16	21	25	29	34	38	42	46	50	55	59	63	67	71	76	80	84	88
7	19	24	29	34	38	43	48	53	58	63	67	72	77	82	86	91	96	101
8	21	27	32	38	43	49	54	60	65	70	76	81	87	92	97	103	108	113
9	23	30	36	42	48	54	60	66	72	78	84	90	96	102	108	114	120	126
10	26	33	39	46	53	60	66	73	79	86	93	99	106	112	119	125	132	138
11	28	36	43	50	58	65	72	79	87	94	101	108	115	122	130	137	144	151
12	31	39	47	55	63	70	78	86	94	102	109	117	125	132	140	148	156	163
13	33	42	50	59	67	76	84	93	101	109	118	126	134	143	151	159	167	176
14	35	45	54	63	72	81	90	99	108	117	126	135	144	153	161	170	179	188
15	38	48	57	67	77	87	96	106	115	125	134	144	153	163	173	182	191	200
16	40	50	61	71	82	92	102	112	122	132	143	153	163	173	183	193	203	213
17	42	53	65	76	86	97	108	119	130	140	151	161	172	183	193	204	214	225
18	45	56	68	80	91	103	114	125	137	148	159	170	182	193	204	215	226	237
19	47	59	72	84	96	108	120	132	144	156	167	179	191	203	214	226	238	250
20	49	62	75	88	101	113	126	138	151	163	176	188	200	213	225	237	250	262

7.1