

İKİ ORTALAMANIN KARŞILAŞTIRILMASINDA ÖRNEK BÜYÜKLÜĞÜNÜN TAHMİNİ VE BAZI İSTATİSTİK PROGRAMLARIN KULLANIMI

Mehmet TOPAL

*Atatürk Üniversitesi, Ziraat Fakültesi, Zootekni Bölümü, Erzurum
e-mail: mtopal@atauni.edu.tr*

Recep ÖZ

Erzincan Üniversitesi, Eğitim Fakültesi, Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü, Erzincan

Geliş Tarihi / Received : 25.09.2007

Özet: Bağımlı ve bağımsız grup ortalamalarının karşılaştırılmasında belirli I. ve II. tip hata seviyelerinde örnek büyüklüğünün en az ne kadar alınması gerektiği üzerinde durulan çalışmada, örnek büyüklüğünü tahminde kullanılan istatistik paket programlardan PS ve PASS programları kullanıldı. Her iki programda da benzer sonuçlar alındı. Bağımlı ve bağımsız grupların ortalamalarının karşılaştırılmasında örnek büyüklüğü belirlenirken I tip hata (α) seviyesi 0,05, 0,025, 0,01 ve 0,001, II. Tip hata (β) seviyesi ise 0,10 ve 0,20 olarak alındı. Hata seviyeleri küçüldükçe örnek büyüklüğü artmaktadır. Ayrıca populasyon ortalamaları arasındaki fark ve örnek değerlere ait standart sapma değerinin büyüklüğü de örnek büyüklüğü üzerinde önemli etkide bulunmaktadır.

Anahtar Kelimeler: Örnek büyüklüğü, hata seviyesi, testin gücü

Estimation of Sample Size in Comparison of Two Means (Paired and Independent) and Use of Some Statistic Programmes

Abstract: In this study, we focused on what the least sample size is at I and II type error level in comparisons of paired and independent group averages and used PS and PASS programmes to estimate sample size. Similar results were obtained from both programmes. In the comparison of the averages of paired and independent groups, we took 0,05, 0,025, 0,01 and 0,001 as I type error (α), 0,10 and 0,20 as II type (β). As error level decreases, so does the size of sample increases. In addition, the difference between population averages and the extent of standard deviation value belonging to sample values have an important effect on sample size.

Key Words: Sample size, error probability, power of test

GİRİŞ

Belli bir alan veya zamanda üzerinde durulan karakterle ilgili elemanlar topluluğuna populasyon, populasyondan şansa bağlı olarak seçilen üzerinde inceleme ve araştırma yapılan elemanlar topluluğuna ise örnek denir. Araştırmalarda populasyonla çalışmak çok zaman alıcı ve masraflı olduğu için genelde örneklerle çalışılır. Populasyondan çekilen örnekten elde edilen istatistiklerin populasyon parametrelerini sapmasız olarak tahmin etmesi için uygulanacak örnekleme metodu ve çekilen örneğin büyüklüğü önemlidir. Yapılan araştırmalarda elde edilen sonuçların populasyona genelleştirilebilmesi için çekilen örnek populasyonu temsil etmelidir. Bunun için çekilen örnekte şansa bağlılığa dikkat edilmeli ve örnek büyüklüğü yeterli olmalıdır. Bir araştırmanın planlama ve uygulamasında temel prensipler; araştırmanın planlanması ve hazırlık çalışmasının yapılması, örnekleme çerçevesinin ve planının belirlenmesi, hipotezlerin kurulması, ön test ve pilot çalışmanın yapılması, verilerin toplanması, istatistik analizi yapılarak kararların verilmesi, sunumu ve sonuçların yayınlanması gibi sıralanabilir. Yapılan bir çalışmada üzerinde çalışılacak örneğin büyüklüğünün tahmini araştırmaların planlanmasında en önemli safhalardan biridir.

Doğru bir hipotezin reddedilmesiyle işlenen hataya I. tip hata (α), yanlış bir hipotezin kabul edilmesi halinde işlenen hataya II. tip hata (β) denir. Bu iki hata tipinden herhangi birinin büyümesi diğerinin küçülmesini sebep olur. Ancak örnek büyüklüğünü artırmakla iki hata tipi birlikte kontrol altında tutulur. Testin gücü, $1-\beta$ ile ölçülür (Yıldız ve Bircan 1994).

Cochran (1963), Sampford (1962) ve Som (1996), örnekleme metotları ve örnek sayısının tahmini üzerine çalışmışlar, Lunsford ve Lunsford (1995a) ise şansa bağlı ve amaçlı örnekleme planlarını açıklamışlardır. Adcock (1997) örnek büyüklüğünün tahmin metotları ile ilgili bir çalışma yapmış ve araştırmasında doğrusal regresyonda örnek büyüklüğünün tahminini anlatmıştır. Guenther (1975) t testi için örnek büyüklüğü formülü geliştirmiş ve bu formülü çeşitli değerler üzerinde uygulamıştır. Örnek büyüklüğünün tahmininde birçok istatistik paket program kullanılmaktadır. Bir veya iki örnek ortalamasının karşı-

laştırılmasında yeterli örnek büyüklüğünü tahminde G*Power, PASS, NQuery Advisor, PC-Size, PS programlarından yararlanılmaktadır.

Çalışmada, bağımlı ve bağımsız grup ortalamalarının karşılaştırılmasında belirli I. ve II. tip hata seviyelerinde optimum örnek büyüklüğünün tahmin edilmesi üzerinde duruldu ve örnek büyüklüğünü tahminde kullanılan istatistik paket programlardan PS ve PASS programları kullanıldı.

MATERYAL VE METOT

Araştırma materyalini, Atatürk Üniversitesi Ziraat Fakültesinde yetiştirilen yumurtacı tavuklar ve broyler piliçler oluşturmaktadır. Bağımsız grupların karşılaştırılması için yumurtacı tavukların iki farklı sıcaklıkta ($20C^{\circ}$ ve $30C^{\circ}$) yem tüketim değerleri, bağımlı gruplarda örnek büyüklüğünün belirlenmesinde ise 20 adet Broyler piliçin iki farklı sıcaklıktaki hematokrit değerleri kullanıldı. Populasyon ortalaması, populasyon değerlerinin merkezi noktasını gösterir. Populasyonların birbirinden farklı olup olmadığı incelenirken populasyon ortalamalarının farklılığına bakılır. Eğer populasyon ortalamaları farklı ise populasyonların farklı olduğu ifade edilir. Karşılaştırılacak iki populasyonun dağılışının normale uyduğu ve varyanslarının homojen olduğu durumlarda populasyonlardan çekilecek örnek ortalamalarının karşılaştırılmasında t ve Z testleri kullanılır. Populasyon varyansı biliniyorsa Z dağılışı, populasyon varyansı bilinmediği zaman t dağılışı kullanılır.

Belirli I. tip (α) ve II. tip (β) hata seviyesinde bağımlı iki grup ortalamaları karşılaştırılırken kullanılacak örnek büyüklüğü;

$$n = \frac{S^2(t_{\alpha/2, n-1} + t_{\beta, n-1})}{d^2} \quad \text{veya} \quad n = \frac{\sigma^2(Z_{\alpha/2} + Z_{\beta})}{d^2}$$

Bağımsız iki grup karşılaştırılırken kullanılacak örnek büyüklüğü;

$$n = \frac{2S^2(t_{\alpha/2, (n1+n2-2)} + t_{\beta, (n1+n2-2)})}{d^2} \quad \text{veya} \quad n = \frac{2\sigma^2(Z_{\alpha/2} + Z_{\beta})}{d^2}$$

eşitlikleri ile tahmin edilir (Lunsford ve Lunsford, 1995b).

d: Populasyon ortalamaları arasındaki fark ($\mu_1 - \mu_2$)

Bir veya iki örnek ortalamasının karşılaştırılmasında yeterli örnek büyüklüğünü tahminde PASS ve PS programları kullanıldı. PS programı <http://biostat.mc.vanderbilt.edu/twiki/bin/view/Main/PowerSampleSize> internet adresinden, PAS programı ise <http://www.ncss.com/pass.html> adresinden indirildi. PS programının kullanım süresi kısıtlı olmamasına rağmen PASS programının yedi günlük kullanım süresi bulunmaktadır.

PS (Power and Sample Size Program) programında iki örnek ortalamasının karşılaştırılmasında örnek büyüklüğü tespit edilirken çıktı (output) kısmında örnek büyüklüğü (Sample Size); Plan (Design) kısmında kullanılacak t testi bağımlı gruplar (paired) veya bağımsız gruplar (Independent) seçilir. Girdi (Input) kısmında ise aşağıdaki değerlerin girilmesi gerekir:

α : İki yönlü test için I. tip hata seviyesi

power : Testin gücü (1- β)

δ : popülasyon ortalamaları arasındaki fark

σ : Bağımsız grupların karşılaştırılmasında gruplar içi standart sapma, bağımlı gruplarda çiftler arasındaki fark değerlere ait standart sapma
m: bağımsız gruplarda iki gruptaki örnek büyüklüğü oranı eğer örnek büyüklükleri eşit ise m=1 olur, bağımlı gruplarda örnek büyüklüğü tespit edilirken m değeri girilmez.

Değerler girildikten sonra hesapla (Calculate) düğmesine tıklanır. Girilen değerlere göre gerekli örnek büyüklüğü hesaplanır.

PASS programında iki örnek ortalamasının karşılaştırılmasında örnek büyüklüğü tespit edilirken bağımsız gruplar için örnek büyüklüğünün tespitinde

Find: örnek büyüklüğünü bulmak için N1 seçilir,

Mean1: 1. grubun ortalaması girilir,

Mean2: 2. grubun ortalaması girilir. HO hipotezinde grup ortalamaları eşit alınmışsa Mean1 ve Mean2 kısmına aynı ortalama değeri girilir veya ortalamalar arasındaki fark girilir,

N1: 1. grubun örnek büyüklüğü,

N2: 2. grubun örnek büyüklüğü,

Alternative Hypothesis: Alternatif hipotez (H1) belirlenir,

Alpha: I. tip hata ihtimal seviyesi girilir,

Beta (1-Power): II. tip hata ihtimal seviyesi girilir,

S1: 1. grubun standart sapması,

S2: 2. grubun standart sapması girilir.

Bağımlı grupların testinde örnek büyüklüğünü tespit etmek için

Find: Örnek büyüklüğünü bulmak için N seçilir

Mean0: Bağımlı grupların karşılaştırılmasında HO hipotezinde fark değerlere ait ortalama sıfır (0) olarak alındığı için Mean0 yerine sıfır değeri girilir

Mean1: Alternatif (H1) hipotezinde belirlenen eşler arasındaki farkların ortalaması.

N: Ön çalışmadaki örnek büyüklüğü

S: Eşler arasındaki fark değerlerinin standart sapması

Alternative Hypothesis: Alternatif hipotez (H1) belirlenir

Alpha: I. tip hata ihtimal seviyesi girilir

Beta (1-Power): II. tip hata ihtimal seviyesi girilir

ARAŞTIRMA BULGULARI VE TARTIŞMA

Bağımsız grupların karşılaştırılması için yumurtacı tavukların iki farklı sıcaklık (20°C ve 30°C) seviyesindeki yem tüketim değerleri alınmıştır. Yumurta tavuklarında sıcaklığın yem tüketimine etkisini incelemek amacıyla 56 yumurta tavuğu iki farklı sıcaklıkta (20°C ve 30°C) denemeye tabi tutularak yumurta tavuklarının 75 gün içerisinde yem tüketimine ait ortalama (\bar{x}), standart sapma (S) ve standart hata (SX) değerleri Çizelge 1'de verilmiştir. Bağımlı gruplarda örnek büyüklüğünün belirlenmesinde 20 adet Broyler piliçler iki farklı sıcaklıktaki hematokrit değerleri kullanıldı. Broyler piliçler önce 23°C'de, bir hafta sonra 16°C'de tutularak hematokrit değerleri ölçülmüş ve sıcaklığın hematokrit değerinden ne yönde nasıl bir etki yaptığı gözlenmeye çalışılmış. Broyler piliçlerin hematokrit değerlerine ait tanıttıcı istatistik değerler Çizelge 3'de verildi.

Çizelge 1: Yumurta tavuklarında yem tüketim ve yumurta verimlerine ait deskriptif değerler

	Sıcaklık C ⁰	N	\bar{X}	S	S _X	P
Yem	20	28	113.89	3.44	0.65	0.001
Tüketimi	30	28	92.17	2.68	0.51	

Populasyon varyanslarının bilinmediği bağımsız iki grup ortalamasının karşılaştırılmasında farklı I. ve II. tip hata seviyelerinde ve populasyon ortalamaları arasındaki fark (d) ve standart sapma (S) değerlerinde PASS ve PS programlarına göre elde edilen sonuçlar Çizelge 2 de verildi. Bağımsız grup ortalamalarının karşılaştırılmasında t dağılışı kullanıldı.

Çizelge 2: Bağımsız grupların karşılaştırılmasında farklı I. ve II. tip hata seviyelerinde örnek büyüklüğü değerleri.

d	β	α	PS			PASS		
			S=3,08	S=1,0	S=0,5	S=3,08	S=1,0	S=0,5
0.2	0.10	0.050	4985	526	5	4856	526	132
		0.025	5888	622	5	5736	621	156
		0.010	7059	745	6	6876	745	186
		0.001	9917	1048	9	9660	1046	262
	0.20	0.050	3724	393	4	3627	393	100
		0.025	4510	476	4	4393	476	121
		0.010	5541	585	5	5397	585	148
		0.001	8101	856	8	7891	855	217
0.5	0.10	0.050	798	85	22	778	86	23
		0.025	943	100	26	918	101	27
		0.010	1131	121	31	1101	121	32
		0.001	1589	170	44	1547	170	45
	0.20	0.050	597	64	17	581	64	17
		0.025	723	77	20	703	78	21
		0.010	888	95	25	864	96	26
		0.001	1298	139	37	1263	140	37
1.0	0.10	0.050	200	22	6	196	23	7
		0.025	237	26	8	231	27	8
		0.010	284	31	9	277	32	10
		0.001	399	44	13	387	45	14
	0.20	0.050	150	17	5	147	17	6
		0.025	182	20	6	177	21	7
		0.010	223	25	8	218	26	8
		0.001	326	37	11	316	37	12
1.5	0.10	0.050	90	10	4	88	11	4
		0.025	106	12	4	104	13	5
		0.010	127	15	5	124	15	6
		0.001	179	21	7	175	22	8
	0.20	0.050	67	8	3	66	9	4
		0.025	81	10	4	80	10	4
		0.010	100	12	4	98	13	5
		0.001	146	18	6	143	18	7

Bağımsız grupların karşılaştırılmasında örnek büyüklüğünü tespit için kullanılan iki programda da

benzer sonuçlar alındı. Her iki programda I ve II tip hata seviyesi küçüldükçe örnek büyüklüğü artmaktadır. Yani testin gücü (1- β) büyüdükçe örnek büyüklüğünde artış olmaktadır. Ayrıca gruplara ait standart sapma değerleri ve ortalamalar arasında olması istenen fark (d) da örnek büyüklüğü üzerinde önemli bir etkide bulunmaktadır. Standart sapma değeri küçüldükçe örnek büyüklüğünde azalma olurken, populasyon ortalamaları arasındaki fark (d) küçüldükçe örnek büyüklüğünde artış olmaktadır. Standart sapma 0,5 ile 1 arasında olduğunda ve populasyon ortalamaları arasındaki fark değeri de 0,5 ile 1 arasında değiştiği durumlarda elde edilen örnek büyüklüğünün daha kullanışlı ve ekonomik olduğu gözlenmektedir. Standart sapma 3 gibi bir değer alındığında ve populasyon ortalamaları arasındaki fark değerinin de çok küçük olması istendiği durumda elde edilecek örnek büyüklüğünün çok büyük olduğu, bununda ekonomik olmayacağı söylenebilir. Çünkü örneklemedeki en önemli avantajlardan biriside daha ekonomik olarak araştırmanın kurulmasıdır.

Çizelge 3: Broyler piliçlerde iki farklı sıcaklıktaki hematokrit değerlerine ait tanıtıcı istatistikler

	Sıcaklık C ⁰	N	\bar{X}	S	$S_{\bar{x}}$	P
Hematokrit	23	20	32,65	1,46	0,33	0.001
	16	20	34,95	1,79	0,40	
Fark		20	2,30	1,49	0,33	

Populasyon varyanslarının bilinmediği bağımlı iki grup ortalamasının karşılaştırılmasında farklı I. ve II. tip hata seviyelerinde ve populasyon ortalamaları arasındaki fark (d) ve eşler arasındaki fark değerlerine ait standart sapma (S) değerlerinde PASS ve PS programlarına göre elde edilen sonuçlar Çizelge 4'de verildi. Bağımlı grup ortalamalarının karşılaştırılmasında t dağılışı kullanıldı. Bağımlı grupların karşılaştırılmasında her iki programa göre çeşitli I. ve II tip hata seviyelerinde elde edilen örnek büyüklükleri Çizelge 4 de verildi. Çizelge 4 incelendiğinde örnek değerler arasındaki veya populasyon değerler arasındaki fark değerlerine ait standart sapma büyüdükçe elde edilen örnek büyüklüğünde de artış olmaktadır.

Çizelge 4: Bağımlı grupların karşılaştırılmasında farklı I. ve II. tip hata seviyelerinde örnek büyüklüğü değerleri.

d	β	α	PS			PASS		
			S=1,49	S=1,0	S=0,5	S=1,49	S=1,0	S=0,5
0.2	0.10	0.050	585	265	68	586	265	68
		0.025	691	313	80	690	313	81
		0.010	829	375	96	827	376	97
		0.001	1165	527	135	1162	529	137
	0.20	0.050	438	198	51	438	199	52
		0.025	530	240	62	531	241	62
		0.010	651	295	76	649	296	77
		0.001	953	432	112	949	433	113
0.5	0.10	0.050	95	44	13	96	44	13
		0.025	113	52	15	113	53	16
		0.010	135	63	18	136	63	19
		0.001	190	88	26	192	90	27
	0.20	0.050	72	33	10	72	34	10
		0.025	87	40	12	87	41	13
		0.010	107	50	15	108	51	16
		0.001	157	73	22	158	74	23
1.0	0.10	0.050	25	13	5	26	13	5
		0.025	30	15	6	31	16	6
		0.010	36	18	7	37	19	8
		0.001	51	26	10	52	27	11
	0.20	0.050	19	10	4	20	10	5
		0.025	24	12	5	24	13	6
		0.010	29	15	6	30	16	7
		0.001	43	22	9	44	23	10
1.5	0.10	0.050	12	7	3	13	7	4
		0.025	15	8	4	15	9	5
		0.010	18	10	5	19	11	6
		0.001	25	14	7	27	15	8
	0.20	0.050	10	6	3	10	6	4
		0.025	12	7	4	13	7	4
		0.010	15	8	4	15	9	5
		0.001	22	12	6	23	13	7

Fark değerler arasındaki ortalama farklılıklar da örnek büyüklüğünü önemli derecede etkilemektedir. Populasyon ortalamaları arasındaki fark değeri küçüldükçe doğal olarak elde edilen örnek büyüklüğü de artmaktadır. Çizelge 4 de verilen değerlere göre en uygun örneklemenin populasyon ortalamaları arasındaki farkın 0,5 olduğunda ve fark değerlere ait standart sapmanın da 1 den küçük olduğu durumda yapılabileceği ifade edilebilir.

Her iki programda da benzer sonuçlar alınmasına rağmen PASS programında örnek sayısının hesaplanmasında değişkenlere ait standart sapmanın yanısıra ortalama değerleri de hesaba katılmaktadır. PASS programında ortalama değerler hesaba katılırken, PS programında populasyon ortalamaları arasında olması istenen fark (d) değeri hesaba katılmaktadır. Her iki testde de I tip hata (α) seviyesi 0,05, 0,025, 0,01 ve 0,001, II. tip hata β seviyesi ise 0,10 ve 0,20 alındı. Genelde örnekleme üzerine yapılan çalışmalarda I. ve II tip hata seviyelerinin bu şekilde alınması önerilmektedir. Ancak I. ve II tip hata seviyeleri yapılacak araştırmanın hassasiyetine göre belirlenebilir. I. ve II tip hata seviyeleri küçüldükçe örnek büyüklüğü artmaktadır. Populasyon ortalamaları arasında olması istenen fark küçüldükçe ve örnek değerlere ait standart sapma büyüdükçe örnek büyüklüğü artmaktadır. Örneklemede en önemli etkenlerden birisi elde edilen örneğin araştırmaya göre ekonomik olması aynı zamanda populasyonu temsil etmesi önemlidir. Bunun için yapılacak örneklemede mümkün olduğu kadar populasyonu temsil eden az sayıdaki örnek üzerinde çalışılmaktadır. Populasyon ortalamaları arasındaki fark ($\mu_1 - \mu_2$), örnek büyüklüğü, I. tip hata seviyesi büyüdükçe ve standart sapma küçüldükçe testin gücü ($1 - \beta$) artmaktadır.

KAYNAKLAR

- Adcock, C. J. (1997). Sample size determination: a review. *The Statistician*, 46 (2), 261-283.
- Cochran, W. G. (1963). *Sampling Techniques*. 2. Baskı. John Wiley & Sons, Inc. 1 - 387.
- Guenther, W. C. (1975). A Sample Size Formula For a Non-Central t Test. *The American Statistician*, 29 (3), 120-121.
- Lunsford, B. R. ve Lunsford, T. R. (1995a). The

Research Sample, Part I: Samling. *Journal of Proshetics and Orthotics*. 7 (3). 105-112.

Lunsford, B. R. ve Lunsford, T. R. (1995b). The Research Sample, Part II: Sample Size. *Journal of Proshetics and Orthotics*. 7 (4). 137-143.

Sampford, M.R. (1962). *An Indroduction to Sampling Theory With Application to Agriculture*. Oliver and Boyd. Edinburg. 206 - 219.

Som, R. K. (1996). *Practical Sampling Techniques*. 2. Baskı. Marcel Dekker, Inc. 1- 485.

PASS.<http://www.ncss.com/pass.html>

PS.<http://biostat.mc.vanderbilt.edu/twiki/bin/view/Main/PowerSampleSize>

Yıldız, N ve Bircan, H. (1994). *Uygulamalı İstatistik*. 4. Baskı, Atatürk Üniversitesi Ziraat Fakültesi Ofset Tesisi, s. 97-106, Erzurum