

Uygun Simulasyon Sayısının Belirlenmesi: Monte Carlo Simülasyon Çalışması

Mehmet MENDEŞ¹

Geliş Tarihi: 04.06.2004

Öz: Biyolojik, sosyal, ziraat, tıp ve ekonomik olayların modellenmesinde simülasyon teknikleri oldukça önemlidir. Dolayısıyla, denemenin planlanması aşamasında kaç simülasyon denemesinin yapılacağı sorusu oldukça önemli bir sorudur. Çünkü, yapılacak parametre tahminlerinin kararlı olup olmaması, simülasyon sayısı ile oldukça ilişkilidir. Bu çalışmada, simülasyon sayısının gerçekleşen 1.Tip hata olasılıklarının kararlılığı üzerine etkisi araştırılmıştır. Çalışmada, 16 farklı simülasyon sayısı bakımından gerçekleşen 1.Tip hata olasılıkları tahmin edilmiştir. Çalışma sonuçları, simülasyon sayısının az olduğu durumlarda aynı deneme koşullarında gerçekleşen 1.Tip hata olasılıkları arasındaki farkın büyüdüğünü göstermiştir. Diğer yandan, çalışma sonuçları örnek hacmi ne olursa olsun genel olarak 50000-70000 simülasyon sayısının uygun bir simülasyon sayısı olduğunu göstermiştir.

Anahtar Kelimeler: Simulasyon, I.Tip hata, örnek hacmi, normal dağılım, varyansların homojenliği

Determining of Suitable Simulation Number: A Monte Carlo Simulation Study

Abstract: Simulation is an important method for modeling biological, social, agricultural, medicine and economic processes. Therefore, at the planning stage of a simulation modeling the question of number of simulation runs (N) is very critical. Because, stabilization of parameter estimation changes depend on number of simulation. This study aimed to investigate effect of number of simulation on stabilize of actual Type I error rate. For this aim, Type I error rates were estimated for 16 different number of simulation conditions. Results of this simulation study suggested that while number of simulation was small, differences among the actual Type I error rates increased. On the other hand, it could be said that 50000-70000 simulation numbers were optimum for all sample sizes.

Key Words: Simulation, Type I error, sample size, normal distribution, homogeneity of variance

Giriş

Matematiksel olarak çözümü çok zor ya da mümkün olmayan problemlerin ampirik olarak çözümünde, yeni geliştirilen testlerin örnekleme dağılımlarının oluşturulmasında, parametre tahminlerinin yapılmasında, aynı amaçla kullanılabilen istatistik testlerinin değişik deneme koşulları altında 1.Tip hata ve testin gücü bakımından karşılaştırılmasında simülasyon tekniklerinden oldukça sık yararlanılır (Olson 1972, Halperin 1976, Fishman ve Moore 1982, Hauck ve Anderson 1984, Morgan 1984, Ripley 1987, Park ve Miller 1988). Monte Carlo simülasyon tekniği, uygulamada en sık kullanılan tekniklerden birisidir (Kleijnen 1974, Chambers 1977, Rubinstein 1981, Bratley ve ark. 1987, Nash 1990, Darlingto 1996, Mooney 1997, Wilcox 2002). Ancak, bu kadar yaygın olarak kullanılan simülasyon tekniklerinden beklenen yararların sağlanabilmesi, yapılan simülasyon sayısı ile oldukça yakından ilişkilidir. Dolayısıyla, deneme başlangıcında uygun simülasyon sayısının belirlenmesi, hem yapılacak tahminlerin kararlı kalması (invariant), hem de elde edilecek sonuçların güvenilirliği bakımından oldukça önemlidir (Kennedy ve Gentle 1980, Karian ve Dudewicz 1991, Diaz-Emparanze 2002). Son yıllarda yapılan simülasyon çalışmalarında her ne kadar simülasyon sayısının 50000 ve 100000 olduğu bazı çalışmalara rastlanılıyorsa da bu tür çalışmaların oldukça sınırlı olduğu görülmektedir (Mendes 2002, Mendes ve Başpınar 2003, Mendes ve Pala 2004). Uygulamada yapılan simülasyon çalışmalarındaki simülasyon sayıları

daha ziyade 1000-30000 arasında değiştiği görülmektedir. Halbuki, 1000 simülasyondan elde edilecek tahminler arasındaki farklarla, 50000 ya da 100000 simülasyondan elde edilecek tahminler arasındaki farklar oldukça büyük olmaktadır. Söz konusu farklar, küçük hacimli örneklerle çalışılması durumunda daha da belirgin olmaktadır. Diğer yandan, teknolojinin oldukça geliştiği ve bilgisayar işlemcilerinin artık Ghz olduğu günümüzde simülasyon sayısının 10000 ya da daha altında tutulması düşündürücüdür. Daha öncede belirtildiği gibi, bir simülasyon denemesi sonucunda yapılan parametre tahminleri, 1.Tip hata olasılıkları ya da testin güç değerlerinin söz konusu simülasyon denemesinin her çalıştırılmasında kararlı kalması (aynı ya da bir birine oldukça yakın) arzu edilir (L'Ecuyer 1988). Bu da, denemedeki simülasyon sayısı ile ilişkilidir. Yani, ampirik olarak gerçekleşen parametre tahminleri, 1.Tip hata ya da testin güç değerleri arasındaki farkların stabil kalması simülasyon sayısına bağlı olarak değişmektedir (Rubinstein 1981, Robey ve Barcikowski 1992, Kromrey ve Hines 1995, Diaz-Emparanze 2002).

Bu çalışmada, farklı sayıdaki simülasyon denemeleri sonucunda gerçekleşen 1.Tip hata olasılıkları tahmin edilmiş ve buradan hareketle yapılacak parametre tahminlerinin ne derecede birbirinden farklılık gösterebileceği üzerinde durulmuştur.

¹ Çanakkale Onsekiz Mart Üniv. Ziraat Fak. Zootekni Bölümü-Çanakkale

Materyal ve Yöntem

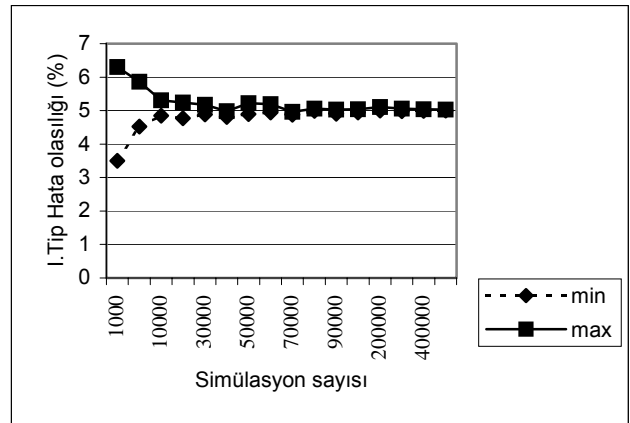
Bu çalışmanın materyalini simülasyon tekniği ile üretilen tesadüf sayıları oluşturmuştur. Tesadüf sayılarının üretilmesinde Anonymus (1999) istatistik paket programının IML kütüphanesinden yararlanılmıştır. Çalışmada, populasyon dağılımlarının normal olduğu durumlar dikkate alınmıştır. Bu amaçla IML kütüphanesinin RNNOR fonksiyonundan yararlanılmıştır. Çalışmada, her birinde 5, 10 ve 20 gözlem bulunan iki grubun karşılaştırıldığı deneme koşulları dikkate alınmıştır. Denemede; 1000, 5000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 200000, 300000, 400000 ve 500000 olmak üzere 16 farklı simülasyon sayısı dikkate alınmıştır. Daha sonra bu 16 farklı koşul için t-testinin gerçekleşen 1.Tip hata olasılıkları elde edilmiştir. Çalışmada, normallik ve varyansların homojenliği ön şartlarının yerine geldiği durumlar dikkate alınmıştır. Bilindiği üzere "gerçekte doğru olan bir hipotezin (H_0) reddedilme olasılığı" 1.Tip hata (α) olasılığı olarak adlandırılır (Bradley 1978, Cohen 1988, Micceri 1989, Sawilowsky ve Blair 1992). Dolayısıyla yapılan N tane simülasyon denemesi sonucunda yanlışlıkla reddedilen H_0 hipotezlerinin sayısının nispi frekansı, ampirik olarak gerçekleşen 1.Tip hata olasılığını verecektir. N; simülasyon sayısı, R; yanlışlıkla reddedilen H_0 hipotezi sayısı ve α ; da gerçekleşen 1.Tip hata olasılığı olmak üzere; dikkate alınan her bir simülasyon denemesi için gerçekleşen 1.Tip hata olasılıkları $\alpha = R / N$ şeklinde hesaplanır.

Bulgular ve Tartışma

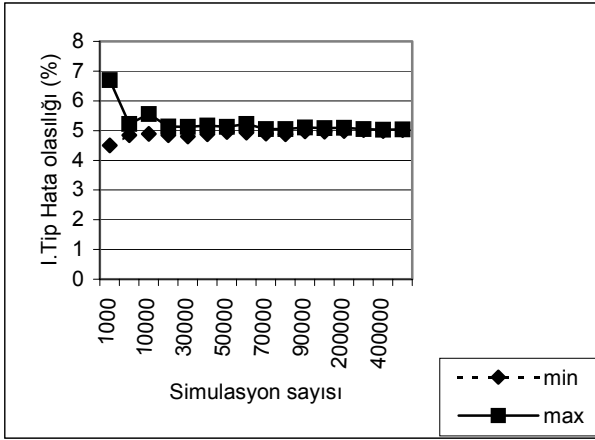
Çalışmada dikkate alınan simülasyon sayılarında gerçekleşen 1.Tip hata olasılıklarının ortalama olarak birbirine oldukça yakın oldukları ve genel olarak hep %5 etrafında toplanma eğiliminde oldukları görülmektedir. Ne var ki, üzerinde durulan özellik bakımından sadece ortalama bakmak, söz konusu örneğin tanıtılması için yeterli olamamaktadır. Dolayısıyla, ortalamanın yanında elde edilen gözlem değerlerinin (bu çalışmada 1.Tip hata olasılıklarının) ortalama ne kadar farklılık gösterdiklerinin de bilinmesi gerekir. Buradan hareketle simülasyon sonuçları incelendiğinde; simülasyon sayısının artmasına paralel olarak, ampirik olarak gerçekleşen 1.Tip hata olasılıklarının da giderek birbirine yaklaşıma eğilimine girdikleri görülmektedir. Simülasyon sayısının küçük olduğu denemelerde gerçekleşen 1.Tip hata olasılıklarının birbirinden oldukça büyük farklılıklar gösterdiği görülmektedir. Bu farklılıklar küçük hacimli örneklerle çalışılması durumunda daha da belirginleşmektedir. Mesela, örnek hacmi 5 iken simülasyon sayısının 1000 olması durumunda gerçekleşen 1.Tip hata olasılıklarının %3.50-6.30 arasında değiştiği görülmektedir. Dikkat edileceği üzere bu deneme koşulunda gerçekleşen 1.Tip hata olasılıkları arasındaki fark oldukça büyüktür (%2.80). Simülasyon sayısının 10000'e çıkartılması durumunda gerçekleşen 1.Tip hata olasılıkları arasındaki fark %0.46'ya , 20000 olması durumunda %0.35'e, 30000 olması durumunda %0.28'e , 100000 olması durumunda %0.10'na ve 500000 olması durumunda ise %0.03'e düştüğü gözlenmiştir. Elde edilen

bulgular, Harwell, (1990), Robey ve Barcikowski (1992), Kromrey ve Hines (1995) ve Diaz-Emparanze (2002) yaptıkları çalışmalarla benzerlik göstermektedir. Diğer taraftan, örnek hacminin artmasına paralel olarak gerçekleşen 1.Tip hata olasılıkları arasındaki farka etkisi daha belirgin olarak ortaya konulabilmektedir. Dikkat edileceği üzere simülasyon sayısının arttırılması gerçekleşen 1.Tip hata olasılıkları arasındaki farkın azalmasına (gerçekleşen 1.Tip hatalarının stabil kalmasına) ve daha güvenilir sonuçların elde edilmesine imkan sağlamaktadır. Bu durum, gerçekleşen 1.Tip hata olasılıkları arasındaki farktan görülebileceği gibi, bunlara ilişkin standart hataların azalmasından da görülebilmektedir. Simülasyon sayılarının gerçekleşen 1.Tip hata olasılıklarının stabil kalıp-kalmaması üzerine etkisi, küçük hacimli örneklerle çalışılması durumunda daha belirgin olmaktadır. Diğer taraftan, örnek hacminin 5 olması durumunda bile, 50000 simülasyonla çalışılması gerçekleşen 1.Tip hata olasılıkları arasındaki farkı belirgin bir şekilde azaltmaktadır. Bu fark, simülasyon sayısı arttıkça daha da azalmakta ve simülasyon sayısının 100000'e ulaşması durumunda hemen hemen tam stabil kalmaktadır. Simülasyon sayılarındaki farklılıkların gerçekleşen 1.Tip hata olasılıkları arasındaki farklara etkilerinin görsel olarak ta değerlendirilebilmesi amacıyla çizilen grafik1-grafik 3 birlikte değerlendirildiğinde, örnek hacmi ne olursa olsun, simülasyon sayısının artmasına paralel olarak gerçekleşen 1.Tip hata olasılıkları arasındaki farkların giderek azaldığı dikkati çekmektedir. Dolayısıyla bu durumlarda daha güvenilir sonuçlar elde edilmektedir. Söz konusu bu durum örnek hacminin artmasına paralel olarak daha da belirginleşmektedir. Grafik1-grafik 3 incelendiğinde; örnek hacminin 5 ya da 10 olması durumunda 70000 simülasyonla, örnek hacminin 20 ve daha fazla olması durumunda ise 50000 simülasyonla çalışılmasının uygulamada yeterli olacağı ileri sürülebilir. Diğer yandan, kullanılan bilgisayarların işlemcilerinin hızlarının her araştırmacı için aynı olamayacağı varsayıldığında 50000 ve daha fazla simülasyonun yapılmasının da elde edilecek sonuçların güvenilirliği bakımından yeterli olabileceği söylenebilir.

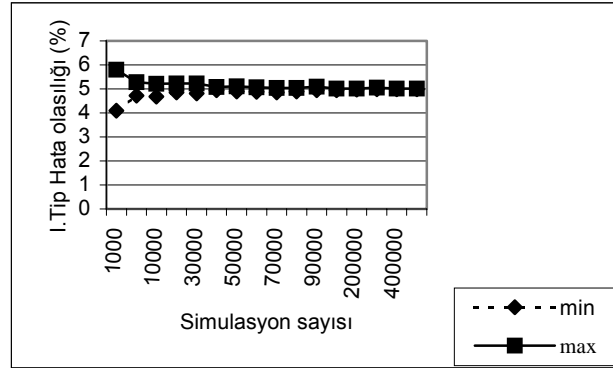
Grafik 1. Simülasyon sayısına bağlı olarak gerçekleşen en küçük ve en büyük 1.Tip hata olasılıklarının dağılımı (n=5)



Grafik 2. Simülasyon sayısına bağlı olarak gerçekleşen en küçük ve en büyük 1.Tip hata olasılıklarının dağılımı (n=10)



Grafik 3. Simülasyon sayısına bağlı olarak gerçekleşen en küçük ve en büyük 1.Tip hata olasılıklarının dağılımı (n=20)



Çizelge 1. Farklı simülasyon sayılarının 1.Tip hata olasılığına etkisi

Sim. sayısı	Var=1:1, n=5				Var=1:1, n=10				Var=1:1, n=20			
	I.Tip hata (%)	Std. hata (%)	I.Tip hata (%)		I.Tip hata (%)	Std. hata (%)	I.Tip hata (%)		I.Tip hata (%)	Std. hata (%)	I.Tip hata (%)	
	\bar{X}	$S_{\bar{X}}$	Min	Max	\bar{X}	$S_{\bar{X}}$	Min	Max	\bar{X}	$S_{\bar{X}}$	Min	Max
1000	5,08	0,39	3,50	6,30	5,48	0,31	4,50	6,70	5,03	0,28	4,10	5,80
5000	5,23	0,24	4,52	5,86	5,06	0,06	4,84	5,22	4,96	0,09	4,72	5,28
10000	5,07	0,08	4,85	5,31	5,20	0,08	4,89	5,56	4,98	0,07	4,68	5,22
20000	5,01	0,07	4,77	5,24	5,00	0,04	4,84	5,14	5,03	0,06	4,86	5,23
30000	5,03	0,04	4,89	5,17	4,96	0,05	4,80	5,12	4,93	0,06	4,82	5,23
40000	4,90	0,02	4,81	4,98	5,05	0,04	4,88	5,17	5,04	0,02	4,95	5,08
50000	5,06	0,04	4,90	5,22	5,05	0,03	4,95	5,12	4,97	0,04	4,89	5,11
60000	5,04	0,04	4,94	5,19	5,04	0,03	4,93	5,22	4,96	0,02	4,88	5,07
70000	4,93	0,01	4,88	4,96	4,99	0,02	4,90	5,05	4,98	0,02	4,86	5,04
80000	5,03	0,01	4,99	5,06	4,99	0,02	4,88	5,05	4,96	0,02	4,89	5,04
90000	4,99	0,02	4,91	5,03	5,06	0,02	4,97	5,10	5,01	0,02	4,94	5,10
100000	5,00	0,01	4,94	5,04	4,98	0,02	4,96	5,08	4,95	0,01	4,95	5,01
200000	5,06	0,01	5,00	5,11	5,02	0,01	4,99	5,09	4,98	0,01	4,97	5,02
300000	5,00	0,01	4,98	5,06	5,03	0,00	5,02	5,05	5,00	0,01	4,98	5,06
400000	5,01	0,01	4,99	5,04	5,00	0,00	4,99	5,03	4,99	0,00	4,99	5,01
500000	5,03	0,00	5,00	5,03	5,02	0,00	5,02	5,04	4,99	0,00	4,99	5,01

Kaynaklar

- Anonymous, 1999. SAS Institute Inc., SAS OnlineDoc®, Version 8, Cary, NC: SAS Institute Inc., Sons.
- Bradley, J. V. 1978. Robustness? British Journal of Mathematical and Statistical Psychology, 31, 144-152.
- Bratley, P., B. L. Fox and L. E. Schrage, 1987. A guide to simulation (2nd ed.). New York: Springer-Verlag.
- Chambers, J. M. 1977. Computational methods for data analysis. New York: John Wiley & Sons.
- Cohen, J. 1988. Statistical Power Analysis for the Behavioral Sciences, 2nd ed. Hillsdale, NJ: Erlbaum.
- Darlington, R. B. 1996. Estimating the true accuracy of regression predictions. Mid-Western Educational Researcher, 9(4), 29-31.
- Diaz-Emparanze, I. 2002. Is small Monte Carlo analysis a good analysis? Chacking the size power and consistency of a simulation based test. Statistical Papers, 43 (4), 567-577.
- Fishman, G. S., L. R. Moore, 1982. A statistical evaluation of multiplicative congruential random number generators with modulus $2^{31}-1$. Journal of the American Statistical Association, 77, 129-136.
- Halperin, S. 1976. Design of Monte Carlo studies. Paper presented at the meeting of the American Educational Research Association, San Francisco, CA. (ERIC Document Reproduction Service No. ED 121 850)
- Harwell, M. R. 1990. Summarizing Monte Carlo results in methodological research. Paper presented at the meeting of the American Educational Research Association, Boston, MA. (ERIC Document Reproduction Service No. ED 319 775)

- Karian, Z. A., E. J. Dudewicz, 1991. Modern statistical systems, and GPSS simulation: The first course. New York Computer Science Press.
- Kennedy, W. J., Jr., J. E. Gentle, 1980. Statistical computing. New York Marcel Dekker.
- Kleijnen, J. P. C. 1974. Statistical techniques in simulation (Part I). New York Marcel Dekker.
- Kromrey, J. D., C. C. Hines, 1995. Use of empirical estimates of shrinkage in multiple regression: A caution. Educational and Psychological Measurement, 55: 901-925.
- L'Ecuyer, P. 1988. Efficient and portable combined random number generators. Communications of the ACM, 31: 742-774.
- Mendes, M. 2002. Normal dağılım ve varyansların homojenliği ön şartlarının gerçekleşmediği durumlarda varyans analizi tekniği yerine kullanılabilir bazı parametrik alternatif testlerin I.tip hata ve testin gücü bakımından irdelenmesi. Ankara Üniv. Fen Bil. Enstitüsü. Doktora Tezi. (Basılmamış)
- Mendes, M., E. Başpınar, 2003. Normal Olmayan Dağılımlı Populasyonlardan Alınan Örneklerde Hesaplanan Çeşitli Test İstatistiklerinin 1.Tip Hata Olasılıkları Bakımından Karşılaştırılması. Ankara Üniv. Ziraat Fak. Tarım Bilimleri Dergisi, 9(1): 23-28.
- Mendes, M., A. Pala, 2004. Evaluation of four tests when normality and homogeneity of variance assumptions are violated. Pakistan Journal of Information and Technology, 4 (1): 38-42.
- Micceri, T. 1989. The unicorn, the normal curve, and other improbable creatures. Psychological Bulletin, 105 (1): 156-166.
- Mooney, C. Z. 1997. Monte Carlo simulation (Sage University Paper series on Quantitative Applications in the Social Sci., series no. 07-116). Thousand Oaks, CA: Sage.
- Morgan, B. J. T. 1984. Elements of simulation. New York Chapman and Hall.
- Nash, J. C. 1990. Compact numerical methods for computers: Linear algebra and function minimisation (2nd ed.). New York Adam Hilger.
- Olson, C. L. 1972. A monte carlo investigation of the robustness of multivariate analysis of variance. Unpublished doctoral dissertation, University of Toronto, Toronto, Canada.
- Park, S. K., K. W. Miller, 1988. Random number generators: Good ones are hard to find. Communications of the ACM, 31: 1192-1201.
- Ripley, B. D. 1987. Stochastic simulation. New York: John Wiley & Sons.
- Robey, R. R., R. S. Barcikowski, 1992. Type I error and the number of iterations in Monte Carlo studies of robustness. British Journal of Mathematical and Statistical Psychology, 45: 283-288.
- Rubinstein, R. Y. 1981. Simulation and the Monte Carlo method. New York John Wiley & Sons.
- Sawilowsky, S. S., R. C. Blair, 1992. A more realistic look at the robustness and type II error properties of the t test to departures from population normality. Psychological Bulletin, 111 (2): 352-360.
- Wilcox, R. R. 2002. Comparing variances of two independent groups, British Journal of Mathematical and Statistical Psychology, 55: 169-175.

İletişim adresi:

Mehmet MENDEŞ

Onsekiz Mart Üniv. Ziraat Fak. Zootekni Bölümü-Çanakkale

Tel: 0 286 218 00 18/1348