Türkiye Klinikleri Biyoistatistik Dergisi
*Turkiye Klinikleri Journal of Biostatistics*
Turkiye Klinikleri J Biostat. 2024;16(3):150-6

ORİJİNAL ARAŞTIRMA ORIGINAL RESEARCH

DOI: 10.5336/biostatic.2024-104312

# Tedavide Başarısızlık Olasılığı İçin Üst Sınır Belirleme Yöntemlerinin Karşılaştırılması: Metodolojik Bir Çalışma

## Comparison of Methods for Setting Upper Limits for Probability of No Treatment Success: A Methodological Study

İsmet DOĞAN[a], Nurhan DOĞAN[a]

[a]Afyonkarahisar Sağlık Bilimleri Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıbbi Bilişim AD, Afyonkarahisar, Türkiye

**ÖZET** Bu çalışmanın amacı, belirli bir tıbbi prosedürün sonucunun olumsuz olma olasılığı p($x$=0) için güven aralığının üst sınırının ($P_U$) tahmin edilmesi ile ilgili önerilen farklı tahmin edicilerin kapsamlı bir karşılaştırmasını sağlamak, %90, %95 ve %99 güven düzeyleri için gereken örneklem büyüklüklerini belirlemektir. **Gereç ve Yöntemler:** $n$ büyüklüğündeki bir örneklemde gözlemlenen başarı sayısı sıfır olduğunda ($x$=0), araştırmacının temel amacı genellikle gerçek başarı olasılığı için makul bir üst sınır elde etmektir. Bu çalışmada $P_U$'nun tahmin edilmesinde Poisson, Clopper-Pearson, Üç Kuralı ve Bayesian Uniform Beta Dağılımı (Mid-p) yöntemlerine dayalı tahmin ediciler ele alınmış, küçük örneklem büyüklükleri ($n\leq50$) için 4 farklı yöntem değerlendirilmiştir. **Bulgular:** $\alpha$ değeri ve/veya örneklem büyüklüğü arttıkça $P_U$ değerleri küçülmektedir. $n\geq45$ için tüm yöntemlerden elde edilen $P_U$ değerleri hemen hemen benzer hâle gelmektedir. Mid-p yönteminden elde edilen $P_U$ değerleri diğer yöntemlerden elde edilen $P_U$ değerlerinden tüm durumlar için daha küçüktür. Poisson yaklaşımı ve Üç Kuralı'ndan elde edilen $P_U$ değerleri birbirlerine oldukça yakındır ve bu yöntemlerden elde edilen $P_U$ değerleri tüm yöntemler içerisinde en büyüktür. Clopper-Pearson yönteminden elde edilen değerler ise arada kalmaktadır. $0{,}01\leq P_U\leq0{,}10$ için örneklem büyüklüğü %99 güvenirlikle $37\leq n\leq461$ kişi, %95 güvenirlikle $22\leq n\leq300$ kişi ve %90 güvenirlikle $15\leq n\leq230$ kişi olmalıdır. **Sonuç:** Çalışmadan elde edilen sonuçlara göre $x$=0 olması durumunda, iyimser bir yaklaşım için Mid-p yönteminin, kötümser bir yaklaşım için ise Üç Kuralı'nın tercih edilmesi gerektiği sonucuna ulaşılmıştır.

**Anahtar kelimeler:** Sıfır tedavi başarısı; Poisson yaklaşımı; Clopper-Pearson yöntemi; Üç Kuralı; Mid-p yöntemi

**ABSTRACT** The aim of this study was to provide a comprehensive comparison of different proposed estimators for estimating the upper limit of the confidence interval ($P_U$) for the probability p($x$=0) of an unfavourable outcome of a given medical procedure, and to determine the sample sizes required for %90, %95 and %99 confidence levels. **Material and Methods:** When the number of observed successes in a sample of size $n$ is zero $x$=0, the main goal of the researcher is usually to obtain a reasonable upper bound for the true probability of success. In this study, estimators based on Poisson, Clopper-Pearson, Rule of Three and Bayesian Uniform Beta Distribution (Mid-p) methods are considered to estimate $P_U$ and four different methods are evaluated for small sample sizes ($n\leq50$). **Results:** $P_U$ values become smaller as $\alpha$ value and/or sample size increases. For $n\geq45$, $P_U$ values obtained from all methods become almost similar. The $P_U$ values obtained from the mid-p method are smaller than the $P_U$ values obtained from the other methods for all cases. The $P_U$ values obtained from Poisson approximation and Rule of Three are very close to each other and the $P_U$ values obtained from these methods are the largest among all methods. The values obtained from the Clopper-Pearson method are in between. For $0.01\leq P_U\leq0.10$, the sample size should be $37\leq n\leq461$ people with 99% confidence, $22\leq n\leq300$ people with 95% confidence and $15\leq n\leq230$ people with 90% confidence. **Conclusion:** According to the results obtained from the study, in case $x$=0, it is concluded that the Mid-p method should be preferred for an optimistic approach and the Rule of Three method should be preferred for a pessimistic approach.

**Keywords:** Zero treatment success; Poisson approximation; Clopper-Pearson method; Rule of three; Mid-p method

**KAYNAK GÖSTERMEK İÇİN:**
Doğan İ, Doğan N. Tedavide başarısızlık olasılığı için üst sınır belirleme yöntemlerinin karşılaştırılması: Metodolojik bir çalışma. Turkiye Klinikleri J Foren Sci Leg Med. 2024;16(3):150-6.

**Correspondence:** Nurhan DOĞAN
Afyonkarahisar Sağlık Bilimleri Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıbbi Bilişim AD, Afyonkarahisar, Türkiye
**E-mail:** nurhandogan@hotmail.com

Peer review under responsibility of Turkiye Klinikleri Journal of Biostatistics.

***Received:*** 06 Jun 2024 ***Received in revised form:*** 04 Oct 2024 ***Accepted:*** 09 Oct 2024 ***Available online:*** 25 Oct 2024

Hekimler sıklıkla hastalarına belirli bir tıbbi prosedürün riskine veya belirli bir sağlık sonucunun olasılığına ilişkin tahminler sunmak zorundadır. Bunu yapmak için literatürde mevcut olanların yanı sıra kendi deneyimlerinden elde ettikleri verileri de kullanabilirler. Genel hedef, tahmini oldukça dar bir aralıkta konumlandırmaktır.[1] Özellikle tıbbi çalışmalarda, başarısız olma riskinin/olasılığının empoze edilen değeri aşmamasını sağlamak, eğer aşıyorsa bunun mümkün olan en küçük değer olduğundan emin olmak önemlidir.[2] Bilinmeyen parametrelerin tahmini yaygın bir istatistiksel problemdir. Örneğin, bir grup hasta bir tür ilaç alıyorsa, hekimler genellikle belirli olumsuz yan etkileri yaşaması beklenen hasta sayısını/oranını bilmek isterler. Bu ilacı daha önce kullanan hastalara ilişkin veriler kullanılarak, söz konusu yan etkileri yaşaması beklenen hasta sayısı/oranı için bir tahmin sağlanabilir. Bu örnekte, belirli koşullar altında tahmin edilmek istenilen parametrenin binom oranı $(p)$ olduğu ve yan etkilere sahip hasta sayısının binom dağılımı ile tanımlanabileceği söylenebilir. Hesaplanan tahmin ya bir nokta tahmini ya da bir aralık tahminidir. Bu örnekte nokta tahmini, yan etkileri olan hasta sayısı, aralık tahmini ise $p$ parametresinin içinde yer alması beklenen bir aralık olacaktır.[3] Matematiksel bir modelle ilgilenmek, verilerin belirli parametrelere uyduğuna dair varsayımlarda bulunmaya dayanır. Binom modeli söz konusu olduğunda varsayımlar; her denemenin iki olası sonucu (başarı ya da başarısızlık) olduğu, her denemenin aynı başarı olasılığına sahip olduğu ve her denemenin karşılıklı olarak birbirini dışladığı veya birbirinden bağımsız olduğudur. Klinik alanlarda ve diğer uygulamalı alanlarda yapılan birçok araştırma, binom dağılımının varsayımlarını karşılamaktadır. Bu çalışmalarda $n$ büyüklüğündeki bir örneklemde gözlenen başarı sayısının $x = 0$ olması nadir karşılaşılan bir durum değildir.[4] Binom dağılımı, $n$ bağımsız deneme sayısını $p$ ilgilenilen olayın olasılığını göstermek üzere iki sonuçlu (başarı veya başarısızlık) denemelerde kullanılan kesikli bir olasılık dağılımıdır. $n$ ve $p$ parametrelerine sahip binom dağılımlı bir $X$ rastgele değişkeninin olasılık fonksiyonu,

$$P(X = x) = \frac{n!}{x!\,(n-x)!}p^x(1-p)^{n-x} \qquad x = 0, 1, 2, \dots\dots\dots\dots, n \qquad (1)$$

eşitliği ile gösterilir.[5] Tıbbi çalışmalardan elde edilen sonuçlar sunulurken güven aralıklarının kullanılması teşvik edilmektedir. Tahmini bir orana ilişkin güven aralığının hesaplanması bu tür uygulamaların basit bir örneğidir ve güven aralığının belirlenmesinde yaygın olarak

$$\hat{p} \mp \sqrt{\frac{\hat{p} - (1-\hat{p})}{n}} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (2)$$

eşitliği kullanılır. $p$ parametresine ilişkin olağan güven aralığının belirlenmesinde, ilgilenilen olaya ilişkin örneklemden elde edilen tahmin oranından $(\hat{p} = x/n)$ yararlanılmaktadır.[6] $x/n$'nin $p$ parametresine ilişkin en çok olabilirlik tahmini olduğu bilinmektedir. Ancak $x = 0$ veya $x = n$ olduğunda, bu tahmin genellikle gerçekçi değildir, alt ve üst sınır belirlemede başarısız olur. Bu gibi durumlarda, en çok olabilirlik tahmini kullanılamaz hâle geldiği için alternatif yöntemler kullanılmalıdır.[7] Dolayısıyla ilk bakışta sorunun çok basit olduğu ve açık ve net bir çözümü olduğu düşünülebilir. Ancak, problem anlamsız karmaşıklıklara sahip zor bir problemdir. Eşitlik 2 ile verilen aralığın gerçek kapsama olasılığının, $p \cong 0$ veya $p \cong 1$ için zayıf olduğu yaygın olarak kabul edilmekte ve genellikle yalnızca $n * \min(p, 1-p) \geq 5$ veya $n * \min\,(p, 1-p) \geq 10$ olduğunda kullanılması gerektiği belirtilmektedir. Ayrıca, Eşitlik 2'de verilen güven aralığı ile ilgili ilgi çekici bir durum da güven aralığının gerçek kapsama olasılığının hem $p$ hem de $n$ değiştikçe ihmal edilemez salınımlar içermesidir. Bazı $(p, n)$ çiftleri için gerçek kapsama olasılığı $C(p, n)$ nominal seviyeye çok yakın veya daha büyük, bazı $(p, n)$ çiftler için ise $C(p, n)$ nominal seviyeden çok daha küçüktür. Salınım hem sabit $p$ için $n$'den hem de sabit $n$ için $p$'den kaynaklanmaktadır. Dolayısıyla kapsama olasılığında $n$ ve $p$ değerlerine bağlı olarak ciddi değişiklikler meydana gelmektedir.[8] Gerçek $p$ sıfıra veya bire yakın olduğunda olağan yaklaşımın zayıf olduğu bilinmektedir. Pires ve Amado tarafından yapılan çalışmada, 20 farklı yöntem ele alınmış ve kapsama olasılığı ile beklenen uzunluk açısından karşılaştırılmıştır.[9] Bir klinisyenin yeni bir prosedürün güvenirliğini kanıtlaması (yani, advers olay olasılığının kabul edilebilir küçük bir olasılıktan daha düşük olduğunu göstermesi) gereken klinik araştırmalarda güvenirlik değerlendirmesi bağlamında, $p$ için alt sınırın pratikte ilgi çekici olmadığını belirtmek gerekir. Klinisyen esas olarak “en kötü durum senaryosunu” veya bir hastayı riske atmak ile ilgili en büyük olasılığı temsil ettiği için $P_U$ ile ilgilenir. Bu nedenle $p$ için alt sınır sıfır olarak alınabilir.[10] Binom dağılımının kesikli doğası nedeniyle, bir üst güven sınırı $(P_U)$ belirlemek sezgisel değildir. $P_U$ için $100(1-\alpha)\%$ güven sınırı tahmin etmeyi amaçlayan çeşitli yöntemler vardır. Ancak bu yöntemlerin çoğu, belirtilen güven düzeylerini koruyamadıkları veya aşırı tutucu oldukları için eleştirilmiştir.[5] Standart yöntemlerin çoğu asimptotik yaklaşımlara dayanmaktadır ve yaklaşımların geçerliliği her zaman belirtilmemekte veya yazardan yazara farklılık göstermektedir.[9] Bu makalenin amacı, $x=0$ durumunda $P_U$’nun tahmin edilmesi için önerilen Poisson, Clopper-Pearson, Üç Kuralı ve Bayesian Uniform Beta Dağılımı (Mid-p) yöntemlerine dayalı tahmin edicilerin kapsamlı bir karşılaştırmasını sağlamak, %90, %95 ve %99 güven düzeyleri için gereken örneklem büyüklüklerini belirlemektir. Çalışmada Helsinki Deklarasyonu prensipleri dikkate alınmıştır.

## GEREÇ VE YÖNTEMLER

$n$ büyüklüğündeki bir örneklemde gözlemlenen başarı sayısı sıfır olduğunda ( $x=0$), araştırmacının temel amacı genellikle gerçek başarı olasılığı için makul bir $P_U$ değeri elde etmektir. Bu amaçla farklı olasılık dağılımları dikkate alınarak üretilmiş birçok çözüm önerilmiştir. Ancak, binom deneyinden elde edilen başarı sayısının sıfır ya da sıfıra çok yakın olduğu durum en çok sorun yaratan durumdur. Bu durumda olabilirlik fonksiyonu giderek daha çarpık hâle gelir ve bu da bazı yaklaşımlara kolayca uyum sağlamaz. Binom dağılımına normal yaklaşım özellikle dikkat çekicidir. Bu yaklaşım, başarı oranı sıfıra veya bire çok yakın olduğunda örneklem büyüklüğünün çok büyük olduğu durumlar dışında kullanılmamalıdır.[11] Bu çalışmada, $x=0$ durumu için $P_U$’nun tahmin edilmesine yönelik binom, Poisson ve beta dağılımını esas alan 4 farklı yöntem dikkate alınmıştır.

### CLOPPER-PEARSON YÖNTEMİ

Clopper-Pearson (1934) tarafından, $P_U$ için kesin (exact) yöntem olarak adlandırılan yöntem önerilmiştir. Yönteme göre $P_U$ için $100(1-\alpha)\%$ sınırı, $x=0$ için,

$$p_U = 1-\alpha^{1/n} \quad (3)$$

eşitliği yardımıyla hesaplanmaktadır.[12]

### Mid-p YÖNTEMİ

Özellikle $n$ değerinin küçük olması durumunda Clopper-Pearson yönteminin olası aşırı tutuculuğunun üstesinden gelmeye yardımcı olmak için Mid-p ayarlamasının kullanılması önerilmektedir. Mid-p yöntemine göre $x=0$ durumunda $P_U$ için yaklaşık 100 (1-α)% güven sınırı,

$$p_U = 1-(2\alpha)^{1/n} \quad (4)$$

eşitliği kullanılarak hesaplanır.[4]

### Poisson Yaklaşımı

Beta dağılımları, binom dağılımları için standart eşlenik önsellerdir ve $p$ üzerinde çıkarım yapmak için beta önsellerini kullanmak oldukça yaygındır. Mid-p olarak isimlendirilen tahmin edici bu yaklaşım kullanılarak elde edilmiştir.[8] $n$ değerinin büyük, $p$ değerinin küçük olduğu durumlarda binom dağılımı, $\lambda = np$ parametreli Poisson Dağılımı’na yakınsar. Dolayısıyla Poisson Dağılımı $p$ için yaklaşık güven sınırları elde etmek amacıyla kullanılabilir. Bu durumda $x=0$ için,

$$p_U = \frac{-\ln(\alpha)}{n} \quad (5)$$

eşitliği kullanılarak $P_U$ için yaklaşık $100(1-\alpha)\%$ güven sınırı hesaplanır.[13]

### Üç Kuralı

%95 güvenirlikle üst sınır için Poisson yaklaşımı,

$$p_U = \frac{3}{n} \dots \dots (6)$$

eşitliği ile verilen "Üç Kuralı"na eş değerdir. Üç Kuralı, ilgilenilen olayın gözlemlenmediği $0/n$ durumları için $3/n$'nin binom parametresi $p$ için yaklaşık %95 güvenirlikle bir üst sınır olduğunu belirtir.[13,14] Sıfır pay problemi, makul olarak mümkün olan ancak mevcut verilerde henüz gerçekleşmemiş bir olayın olasılığının tahmin edildiği durumdur.[15] $3/n$'nin önemli bir matematiksel temeli de bulunmaktadır. Yaklaşık olarak $\alpha^{1/n}$'nin birinci dereceden Taylor serisi açılımına eşittir.[16] Eşitlik 6'nın payında yer alan 3 değerinin yerine, %90 güvenirlik için 2,3 değeri, %99 güvenirlik için ise 4,6 değeri kullanılmalıdır.

### $P_U$ İçin Örneklem Büyüklüğünün Tahmini

Bir klinisyen hiçbir olay meydana gelmediği takdirde $p < p_U$ olduğunu %95 güvenirlikle söyleyebilmek için gereken hasta sayısını belirlemek amacıyla bir örneklem büyüklüğü hesaplamasına ihtiyaç duyabilir. Bu durumda,

Eşitlik 3'te uygun düzenlemeler yapıldığında Clopper-Pearson yöntemi için,

$$n = int \frac{\ln(\alpha)}{\ln(1 - P_U)} \dots \dots (7)$$

Eşitlik 4'te uygun düzenlemeler yapıldığında Mid-p yöntemi için,

$$n = int \frac{\ln(2\alpha)}{\ln(1 - P_U)} \dots \dots (8)$$

Eşitlik 5'te uygun düzenlemeler yapıldığında Poisson yaklaşımı için,

$$n = int\{-ln(\alpha^{1/P_U})\} \dots \dots (9)$$

Eşitlik 6'da uygun düzenlemeler yapıldığında Üç Kuralı yaklaşımı için,

$$n = int\left(\frac{3}{P_U}\right) \dots \dots (10)$$

en küçük örneklem büyüklüğü olacaktır.

## BULGULAR

Farklı örneklem büyüklükleri ve $\alpha$ değerleri için çalışmada dikkate alınan yöntemlere ait elde edilen $P_U$ değerleri Tablo 1'de verilmiştir.

Tablo 1'den de görüldüğü üzere $\alpha$ değeri ve/veya örneklem büyüklüğü arttıkça $P_U$ değerleri küçülmektedir. $n \geq 45$ için tüm yöntemlerden elde edilen $P_U$ değerleri hemen hemen benzer hâle gelmektedir. Mid-p yönteminden elde edilen $P_U$ değerleri diğer yöntemlerden elde edilen $P_U$ değerlerinden tüm durumlar için daha küçüktür. Poisson yaklaşımı ve Üç Kuralı'ndan elde edilen $P_U$ değerleri birbirlerine oldukça yakındır ve bu yöntemlerden elde edilen $P_U$ değerleri tüm yöntemler içerisinde en büyüktür. Clopper-Pearson yönteminden elde edilen değerler ise arada kalmaktadır. Farklı $\alpha$ ve $P_U$ değerleri için örneklem büyüklükleri hesaplanmış ve Tablo 2'de verilmiştir.

Tablo 2'den de görüldüğü üzere $x = 0$ durumunda, $0{,}01 \leq P_U \leq 0.10$ için örneklem büyüklüğü %99 güvenirlikle $37 \leq n \leq 461$ kişi, %95 güvenirlikle $22 \leq n \leq 300$ kişi ve %90 güvenirlikle $15 \leq n \leq 230$ kişi olmalıdır. Mid-p yönteminden elde edilen $n$ değerleri diğer yöntemlerden elde edilen $n$ değerlerinden tüm durumlar için daha küçüktür. Poisson yaklaşımı ve Üç Kuralı'ndan elde edilen $n$ değerleri birbirlerine oldukça yakındır ve bu yöntemlerden elde edilen $n$ değerleri tüm yöntemler içerisinde en büyüktür. Clopper-Pearson yönteminden elde edilen değerler ise arada kalmaktadır.

**TABLO 1:** Farklı örneklem büyüklükleri için üst sınır değerleri.

| $n$ | $x = 0$ için $p_U$ değerleri | | | |
|---|---|---|---|---|
| | **Clopper-Pearson** | **Mid-p** | **Poisson Yaklaşımı** | **Üç Kuralı** |
| | $\alpha = 0{,}01$ | | | |
| 5 | 0,60 | 0,54 | 0,92 | 0,92 |
| 10 | 0,37 | 0,32 | 0,46 | 0,46 |
| 15 | 0,26 | 0,23 | 0,31 | 0,31 |
| 20 | 0,21 | 0,18 | 0,23 | 0,23 |
| 25 | 0,17 | 0,14 | 0,18 | 0,18 |
| 30 | 0,14 | 0,12 | 0,15 | 0,15 |
| 35 | 0,12 | 0,11 | 0,13 | 0,13 |
| 40 | 0,11 | 0,09 | 0,12 | 0,12 |
| 45 | 0,10 | 0,08 | 0,10 | 0,10 |
| 50 | 0,09 | 0,08 | 0,09 | 0,09 |
| | $\alpha = 0{,}05$ | | | |
| 5 | 0,45 | 0,37 | 0,60 | 0,60 |
| 10 | 0,26 | 0,21 | 0,30 | 0,30 |
| 15 | 0,18 | 0,14 | 0,20 | 0,20 |
| 20 | 0,14 | 0,11 | 0,15 | 0,15 |
| 25 | 0,11 | 0,09 | 0,12 | 0,12 |
| 30 | 0,10 | 0,07 | 0,10 | 0,10 |
| 35 | 0,08 | 0,06 | 0,09 | 0,09 |
| 40 | 0,07 | 0,06 | 0,07 | 0,08 |
| 45 | 0,06 | 0,05 | 0,07 | 0,07 |
| 50 | 0,06 | 0,05 | 0,06 | 0,06 |
| | $\alpha = 0{,}10$ | | | |
| 5 | 0,37 | 0,28 | 0,46 | 0,46 |
| 10 | 0,21 | 0,15 | 0,23 | 0,23 |
| 15 | 0,14 | 0,10 | 0,15 | 0,15 |
| 20 | 0,11 | 0,08 | 0,12 | 0,12 |
| 25 | 0,09 | 0,06 | 0,09 | 0,09 |
| 30 | 0,07 | 0,05 | 0,08 | 0,08 |
| 35 | 0,06 | 0,04 | 0,07 | 0,07 |
| 40 | 0,06 | 0,04 | 0,06 | 0,06 |
| 45 | 0,05 | 0,04 | 0,05 | 0,05 |
| 50 | 0,05 | 0,03 | 0,05 | 0,05 |

**TABLO 2:** Farklı $\alpha$ ve $P_U$ değerleri için örneklem büyüklükleri.

| $\alpha$ | $P_U$ | Örneklem Büyüklüğü | | | |
|---|---|---|---|---|---|
| | | **Clopper-Pearson** | **Mid-p** | **Poisson Yaklaşımı** | **Üç Kuralı** |
| 0,01 | 0,01 | 458 | 389 | 461 | 460 |
| | 0,05 | 90 | 76 | 92 | 92 |
| | 0,10 | 44 | 37 | 46 | 46 |
| 0,05 | 0,01 | 298 | 229 | 300 | 300 |
| | 0,05 | 58 | 45 | 60 | 60 |
| | 0,10 | 28 | 22 | 30 | 30 |
| 0,10 | 0,01 | 229 | 160 | 230 | 230 |
| | 0,05 | 45 | 31 | 46 | 46 |
| | 0,10 | 22 | 15 | 23 | 23 |

## TARTIŞMA

$0/n$ oranının olduğu durumlar için çıkarım yaparken, gerçek veya uzun dönemli riskin sıfır ile bazı üst sınırlar arasında olduğu belirli bir güvenirlikle söylenmek istenir. Ancak hangi yöntemin kullanılacağı sorusu için kesin bir cevap söz konusu değildir. Hanley ve Lippman-Hand, $n$ büyüklüğündeki bir örneklemde $x = 0$ için "Üç Kuralı" adını verdikleri yöntemin kullanılmasını önermişlerdir.[1] Kural, binom dağılımının Poisson ya da normal dağılıma yakınsaması gibi herhangi bir yaklaşıma bağlı değildir ve her $p$ ve $n$ için güven seviyesinin üzerinde gerçek kapsama sahiptir. Ancak bu yöntemin dezavantajlarından bazıları, hesaplamaların diğer yöntemlere göre daha karmaşık olması, düzenli olarak gereksiz genişlikte aralıklar üretmesi ve kapsama alanının güven düzeyinden çok daha büyük olmasından dolayı fazla tutucu olarak değerlendirilmesidir.[3] McCracken ve Looney tarafından ise birçok popüler yöntemin $x = 0$ olduğu durumda kötü performans gösterdiği belirtilmiş, Clopper-Pearson yönteminin kullanılması önerilmiştir.[4] Reed tarafından, Clopper-Pearson yönteminin, hem kapsama alanının aşım hem de sıfır genişlik aralıklarını ortadan kaldırdığı için kesin yöntem olarak kabul edilmesine rağmen en tutucu yöntem olduğu belirtilmiştir.[17] Clopper-Pearson aralığı hem küçük hem de büyük örnekler için tutucudur ve her zaman olması gerekenden daha geniştir.[18] Agresti ve Gottard'a göre Mid-p yöntemine dayalı çıkarım, kesikli veriler için kullanışlı ve genel amaçlıdır. Mid-p yöntemi kullanılarak elde edilen güven aralığı, kapsama olasılığının tam olarak $1 - \alpha$ veya en az $1 - \alpha$'ya eşit olmasını garanti etmez. Ancak Mid-p yöntemi, kesin yöntemlerin tutucu etkilerini azaltır ve daha kolay anlaşılabilir sonuçlar sağlar. Mid-p yönteminde $x = 0$ olduğunda alt sınır değeri sıfıra, $x = n$ olduğunda üst sınır değeri 1'e eşit olur.[19] Vollset'de Clopper-Pearson ve Mid-p yöntemlerini önermekte ve bu yöntemlerin her zaman güvenle kullanılabileceğini belirtmektedir.[6] Cohen ve Yang'a göre Mid-p yönteminden elde edilen aralıklarının daha kısa uzunlukları ve yeterli kapsama özellikleri, bunları uygulamalı istatistikçi için mantıklı araçlar hâline getirmektedir.[20] Çalışmadan elde edilen sonuçlar, Vollset, Cohen ve Yang ile Agresti ve Gottard'ın önerilerini desteklemektedir.[6,19,20]

## SONUÇ

Çalışmadan elde edilen sonuçlara göre $x = 0$ olması durumunda, en düşük $P_U$ değerlerinin Mid-p yöntemi tarafından, en yüksek $P_U$ değerlerinin ise Poisson yaklaşımı ve Üç Kuralı yöntemi tarafından tahmin edildiği, dolayısıyla iyimser bir yaklaşım için Mid-p yönteminin, kötümser bir yaklaşım için ise Üç Kuralı'nın tercih edilmesi gerektiği sonucuna ulaşılmıştır.

***Finansal Kaynak***

*Bu çalışma sırasında, yapılan araştırma konusu ile ilgili doğrudan bağlantısı bulunan herhangi bir ilaç firmasından, tıbbi alet, gereç ve malzeme sağlayan ve/veya üreten bir firma veya herhangi bir ticari firmadan, çalışmanın değerlendirme sürecinde, çalışma ile ilgili verilecek kararı olumsuz etkileyebilecek maddi ve/veya manevi herhangi bir destek alınmamıştır.*

***Çıkar Çatışması***

*Bu çalışma ile ilgili olarak yazarların ve/veya aile bireylerinin çıkar çatışması potansiyeli olabilecek bilimsel ve tıbbi komite üyeliği veya üyeleri ile ilişkisi, danışmanlık, bilirkişilik, herhangi bir firmada çalışma durumu, hissedarlık ve benzer durumları yoktur.*

***Yazar Katkıları***

*Bu çalışma hazırlanırken tüm yazarlar eşit katkı sağlamıştır.*

## KAYNAKLAR


1. Hanley JA, Lippman-Hand A. If nothing goes wrong, is everything all right? Interpreting zero numerators. JAMA. 1983;249(13):1743-5. [Crossref] [PubMed]
2. Jäntschi L. Binomial distributed data confidence interval calculation: formulas, algorithms and examples. Symmetry. 2022;14:1104. [Crossref]
3. Tollefsen YS. Different approaches for calculating the confidence intervals for the binomial proportion [Bachelor's thesis]. Trondheim: Norwegian University of Science and Technology; 2023. [Cited: May 23, 2024]. Available from: [Link]
4. McCracken CE, Looney SW. On finding the upper confidence limit for a binomial proportion when zero successes are observed. J Biom Biostat. 2017;8(2):338. [Link]
5. İnal C, Günay S. Olasılık ve Matematiksel İstatistik. 2. Baskı. Ankara: Hacettepe Üniversitesi Fen Fakültesi Basımevi; 1982. p.141-6.
6. Vollset SE. Confidence intervals for a binomial proportion. Stat Med. 1993;12(9):809-24. Erratum in: Stat Med 1995;14(8):875. [Crossref] [PubMed]
7. Razzaghi M. On the estimation of binomial success probability with zero occurrence in sample. JMASM. 2002;1(2):326-32. [Crossref]
8. Brown LD, Cai TT, DasGupta A. Interval estimation for a binomial proportion. Stat Sci. 2001;16(2):101-17. [Crossref]
9. Pires AM, Amado C. Interval estimators for a binomial proportion: comparison of twenty methods. REVSTAT. 2008;6(2):165-97. [Link]
10. Jovanovic BD, Levy PS. A look at the rule of three. Am Stat. 1997;51(2):137-9. [Crossref]
11. Ahlers Z. Estimating the necessary sample size for a binomial proportion confidence interval with low success probabilities [MSc thesis]. Manhattan, Kansas: Kansas State University; 2017. [Cited: May 27, 2024]. Available from: [Link]
12. Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika. 1934;26(4):404-13. [Crossref]
13. Leemis LM, Trivedi KS. A comparison of approximate interval estimators for the Bernoulli parameter. Am Stat. 1996;50(1):63-8. [Crossref]
14. Tuyl F, Gerlach R, Mengersen K. The rule of three, its variants and extensions. Int Stat Rev. 2009;77(2):266-75. [Crossref]
15. Winkler RL, Smith JE, Fryback DG. The role of informative priors in zero-numerator problems: being conservative versus being candid. Am Stat. 2002;56(1):1-4. [Crossref]
16. Louis TA. Confidence intervals for a binomial parameter after observing no successes. Am Stat. 1981;35(3):154. [Crossref]
17. Reed JF III. Better binomial confidence intervals. J Mod Appl Stat Methods. 2007;6(1):153-61. [Crossref]
18. Orawo LA. Confidence intervals for the binomial proportion: a comparison of four methods. Open J Stat. 2021;11(5):806-16. [Crossref]
19. Agresti A, Gottard A. Comment: randomized confidence intervals and the Mid-P approach. Stat Sci. 2005;20(4):367-71. [Crossref]
20. Cohen GR, Yang SY. Mid-P confidence intervals for the Poisson expectation. Stat Med. 1994;13(21):2189-203. [Crossref] [PubMed]