

MIT OpenCourseWare
<http://ocw.mit.edu>

14.30 Ekonomide İstatistiksel Yöntemlere Giriş
Bahar 2009

Bu materyale atıfta bulunmak ve kullanım koşulları için <http://ocw.mit.edu/terms> sayfasını ziyaret ediniz.

14.30 Ekonomide İstatistiksel Yöntemlere Giriş Ders Notları 23

Konrad Menzel

12 Mayıs 2009

1. Örnekler

Örnek 1. Varsayalım ki doğum sırasında bebeklerin ağırlığı (pound cinsinden) $X \sim N(7, 1)$ 'e göre dağılmaktadır. Diyelim ki eğer bir doğum uzmanı bebek bekleyen bir anneye zayıf bir diyet önerisinde bulunsaydı, bu öneri bebeğin ortalamadan 1 pound daha hafif (fakat aynı varyansa sahip) doğmasına sebep olurdu. Canlı doğan 10 kişilik bir örneklem için, $\bar{X}_{10} = 6.2$ 'yi gözlemleriz.

- Doğum uzmanı kötü öneride bulunmuyor boş hipotezine karşı kötü öneride bulunuyor alternatif hipotezi için %5'lik bir testi nasıl oluştururuz? Elimizde

$$H_0 : \mu = 7 \text{ 'ye karşı } H_A : \mu = 6$$

var.

Normal dağılım için, bu basit testi sadece örneklem ortalamasına, \bar{X}_{10} , dayandırmanın daha optimal olduğunu göstermiştik, yani $T(x) = \bar{x}_{10}$. H_0 altında, $\bar{X}_{10} \sim N(7, 0.1)$ ve H_A altında $\bar{X}_{10} \sim N(6, 0.1)$ 'dir. Eğer $\bar{X}_{10} < k$ ise test ret eder. Bu nedenle test büyüklüğü %5 olacak şekilde k 'yi seçmeliyiz, yani

$$0.05 = P(\bar{X}_{10} < k | \mu = 7) = \Phi\left(\frac{k - 7}{\sqrt{0.1}}\right)$$

burada $\Phi(\cdot)$ standart normal c.d.f.dir. Bu nedenle aşağıdaki denklemin tersini alarak k 'yi elde ederiz

$$k = 7 + \sqrt{0.01}\Phi^{-1}(0.05) \approx 7 - \frac{1.645}{\sqrt{10}} \approx 6.48$$

Bundan ötürü, $\bar{X}_{10} = 6.2 < 6.48 = k$ olduğu için ret ederiz.

- Bu testin gücü nedir?

$$P(X_{10} < 6.48 | \mu = 6) = \Phi\left(\frac{6.48 - 6}{\sqrt{0.1}}\right) \approx \Phi(1.518) \approx 93.55\%$$

- Varsayalım ki gücü en az %99 olan bir test istiyoruz, gözlemlemek zorunda olduğumuz yeni doğan bebek sayısı n en az ne kadar olmalıdır? n ile değişecek tek şey örneklem varyansdır, bu nedenle bu örneğin birinci bölümünden, kritik değer $k_n = 7 - \frac{1.645}{\sqrt{n}}$ olduğunu buluruz, diğer taraftan \bar{X}_n 'e dayalı testin gücü ve kritik değer k_n aşağıdaki ile verilir:

$$1 - \beta = P(\bar{X}_n < k_n | \mu = 6) = \Phi(\sqrt{n} - 1.645)$$

$1 - \beta \geq 0.99$ olarak ayarlayınca şu koşullu elde ederiz:

$$\sqrt{n} - 1.645 \geq \Phi^{-1}(0.99) = 2.326 \Leftrightarrow n \geq 3.971^2 \approx 15.77$$

Bu tür güç hesaplamaları genellikle bir istatistiki deney veya bir anket çalışması planlandığında yapılır – örneğin, belli bir ilacın büyüklüğünün etkisini araştırmak için bir ilaç testinde kaç tane hasta kullanacağımızı belirlemek gibi. Çok sayıda kişi üzerinde çalışmak veya anket yapmak çoğu zaman maliyetlidir, bu nedenle yeterince büyük bir olasılıkla anlamlı değişiklikleri bulabilmek için bir deneyin büyüklüğünün ne olması gerektiğini önceden bilmek isteriz.

Örnek 2. Varsayalım ki önceki örneğin kurulumuna olduğu gibi sahibiz, fakat varyansı bilmiyoruz. Onun yerine, $S^2 = 1.5$ gibi bir tahminimiz var. Testi nasıl yapardınız? Daha önce tartıştığımız gibi,

$$T := \frac{X_n - \mu_0}{S/\sqrt{n}} \sim t_{n-1}$$

istatistiği, eğer gerçek ortalama μ_0 ise $n - 1$ serbestlik dereceli bir öğrenci t -dağılımıdır. Dolayısıyla eğer aşağıdaki koşul sağlanırsa, H_0 'ı ret ederiz:

$$T = \frac{\bar{X}_n - 7}{S/\sqrt{10}} < t_9(5\%)$$

Problemde verilen rakamları yerine koyarsak, $T = -\frac{0.8}{\sqrt{1.5/10}} \approx -2.066$ olur. Bu da $t_9(0.05) = -1.83$ 'ten küçüktür.

Örnek 3. $X_i \sim \text{Bernoulli}(p)$, $i = 1, 2, 3$, olsun. Yamuk bir parayı birbirinden bağımsız olarak üç kere fırlatıyoruz ve eğer tura gelirse $X_i = 1$ 'dir, diğer durumda $X_i = 0$ 'dir. $H_A : p = 2/3$ 'e

karşı $H_0 : p = 1/3$ 'ü test etmek istiyoruz. Her iki test basit olduğu için, olabilirlik oran testini kullanabiliriz,

$$T = \frac{f_0(X)}{f_A(X)} = \frac{\prod_{i=1}^3 \left(\frac{1}{3}\right)^{X_i} \left(\frac{2}{3}\right)^{1-X_i}}{\prod_{i=1}^3 \left(\frac{2}{3}\right)^{X_i} \left(\frac{1}{3}\right)^{1-X_i}} = \frac{2^{3-\sum_{i=1}^3 X_i}}{2^{\sum_{i=1}^3 X_i}} = 2^{3-2\sum_{i=1}^3 X_i}$$

Bu nedenle eğer aşağıdaki koşul gerçekleşirse ret ederiz.

$$2^{3-2\sum_{i=1}^3 X_i} \leq k \Leftrightarrow (3 - 2\sum_{i=1}^3 X_i) \log 2 \leq \log k$$

burada $\bar{X}_3 \geq \frac{1}{2} - \frac{\log k}{6 \log 2}$ 'ye eşittir. k 'yi belirlemek için, H_0 ve H_A altında \bar{X}_3 'in olası bütün değerlerini ve olasılıklarını listeleyelim:

\bar{X}_3	H_0 altında olasılık	H_A altında olasılık	H_0 altında birikimli olasılık
1	$\frac{1}{27}$	$\frac{8}{27}$	$\frac{1}{27}$
$\frac{2}{3}$	$\frac{6}{27}$	$\frac{12}{27}$	$\frac{7}{27}$
$\frac{1}{3}$	$\frac{12}{27}$	$\frac{6}{27}$	$\frac{19}{27}$
0	$\frac{8}{27}$	$\frac{1}{27}$	1

Böylece eğer testin büyüklüğünün $\alpha = 1/27$ 'ye eşit olmasını arzuluyorsak, sadece ve sadece $\bar{X}_3 > 2/3$ ise ret edebilirdik. Aynı sonucu doğuracak şekilde $k = 2/3$ 'ü seçebiliriz. Bu testin gücü şuna eşittir:

$$1 - \beta = P(\bar{X}_3 = 1|H_A) = \frac{8}{27} \approx 29.63\%$$

Örnek 4. Varsayalım ki aşağıdaki fonksiyon tarafından türetilmiş bir tek gözlemimiz var,

$$f_0(x) = \begin{cases} 2x & \text{eğer } 0 \leq x \leq 1 \text{ ise} \\ 0 & \text{diğer bütün durumlarda} \end{cases} \text{ ya da } f_A(x) = \begin{cases} 2 - 2x & \text{eğer } 0 \leq x \leq 1 \text{ ise} \\ 0 & \text{diğer bütün durumlarda} \end{cases}$$

- $\alpha + \beta$ toplamını minimize eden test prosedürünü bulunuz – eğer $X = 0.6$ ise ret eder miyiz? Sadece bir X gözlemimiz olduğu için, X cinsinden kritik bölgeyi oluşturmak çok karmaşık değildir, bazı ileri düzeyde istatistikleri bulmaya çalışmak çok şey kazandırmayacaktır (ancak Neyman-Pearson burada işe yarayabilir). Yoğunluk grafiğine bakarak, k kritik değerlerinde küçük X değerleri

için testin ret etmesi gerektiği konusunda ikna olabiliriz. Tip I ve Tip II'nin olasılıkları, sırasıyla, $0 \leq k \leq 1$ için şöyledir,

$$\alpha(k) = P(\text{ret} | H_0) = \int_0^k 2x dx = k^2$$

ve

$$\beta(k) = P(\text{ret etme} | H_A) = \int_k^1 (2 - 2x) dx = 2(1 - k) - 1 + k^2 = 1 - k(2 - k)$$

Bu nedenle, k üzerinden hata olasılıklarını minimize ederiz.

$$\min_k \{\alpha(k) + \beta(k)\} = \min_k \{k^2 + 1 - k(2 - k)\} = \min_k \{2k^2 + 1 - 2k\}$$

Minimize edilmiş terimin türevini alıp sıfıra eşitlersek,

$$0 = 4k - 2 \Leftrightarrow k = \frac{1}{2}$$

Dolayısıyla, eğer $X < 1/2$ ise ret etmeliyiz ve $\alpha = \beta = 1/4$ 'tür. Ancak, $X = 0.6$ için H_0 'ı özellikle ret etmiyoruz.

- Bütün testler arasında $\alpha \leq 0.1$ gibi, en küçük β değerli testi bul. β nedir? $X = 0.4$ olsa ret eder miydiniz? – önce k için $\alpha(k) = 0.1$ 'i çözeriz. Yukarıdaki formülü kullanarak, $k = \sqrt{0.1}$ olur. Dolayısıyla,

$$\beta(\bar{k}) = 1 - 2\bar{k} + \bar{k}^2 = 1.1 - 2\sqrt{0.1} \approx 46.75\%$$

$k = \sqrt{0.1} \approx 0.316 < 0.4$ olduğu için, $X = 0.4$ için H_0 'ı ret etmeyiz.

Örnek 5. X geçişkeni $X_i \sim U[0, \theta]$ dağılımlıdır ve varsayalım ki bir i.i.d. örneklem X_1, \dots, X_n 'i gözlemledik ve aşağıdakini test etmek istiyoruz

$$H_0 : \theta = \theta_0$$

$$H_A : \theta \neq \theta_0, \theta > 0$$

İki seçeneğimiz var: θ için bir $1 - \alpha$ 'lık güven aralığını oluşturabiliriz ve eğer θ_0 'ı kapsamazsa ret ederiz. Diğer bir seçenek olarak, bir GLRT testi oluşturabiliriz

$$T = \frac{L(\theta_0)}{\max_{\theta \in \mathbb{R}_+} L(\theta)}$$

Olabilirlik fonksiyonu aşağıdaki ile verilir:

$$L(\theta) = \prod_{i=1}^n f_X(X_i|\theta) = \begin{cases} \left(\frac{1}{\theta}\right)^n & \text{eğer } 0 \leq X_i \leq \theta, i = 1, \dots, n \text{ ise} \\ 0 & \text{diğer bütün durumlarda} \end{cases}$$

T'nin payı maksimize edici üzerinden hesaplanan olabilirlik ile elde edilir. Bu maksimum olabilirlik tahmin edicidir, $\hat{\theta}_{MLE} = X_{(n)} = \max\{X_1, \dots, X_n\}$, yani

$$\max_{\theta \in \mathbb{R}_+} L(\theta) = L(\hat{\theta}_{MLE}) = \left(\frac{1}{X_{(n)}}\right)^n$$

Dolaysıyla

$$T = \frac{L(\theta_0)}{\max_{\theta \in \mathbb{R}_+} L(\theta)} = \left(\frac{X_{(n)}}{\theta_0}\right)^n$$

Test büyüklüğünü arzulanan düzeye eşitleyen istatistiğin kritik değeri k'yi bulmak için, $\theta = \theta_0$ boş hipotezi altındaki dağılımı bilmek zorundayız – bunun için sıra istatistiği bölümüne bakmamız gerekebilir.

Bir dip not olarak, daha önce büyük n'ler için, boş hipotez altında GLRT'nin bir ki-kare, χ^2 , dağılımı olduğunu söylediğimiz halde, bu örnek için bunun doğru olmadığı anlaşılıyor çünkü gerçek parametre değerinde yoğunluk sürekli değildir.

2. Diğer Özel Testler

Varsayalım ki i.i.d. özeliğine sahip iki örneklemimiz var, bunlar X_1, \dots, X_n ve Z_1, \dots, Z_n 'dir ve potansiyel olarak ikisi farklı büyüklüktedir, n_1 ve n_2 gibi. İki farklı dağılım oluşturabiliriz:

$$\begin{aligned} X_i &\sim N(\mu_X, \sigma_X^2) \\ Z_i &\sim N(\mu_Z, \sigma_Z^2) \end{aligned}$$

Yapmaya çalıştığımız iki farklı test şunlardır:

1. $H_0 : \mu_X = \mu_Z$, $H_A : \mu_X \neq \mu_Z$, ya da
2. $H'_0 : \sigma_X^2 = \sigma_Z^2$, $H'_A : \sigma_X^2 \neq \sigma_Z^2$

Bu hipotezleri nasıl test ederiz?

1. Burada sadece σ_X^2 ve σ_Z^2 değerleri bilinen durumu ele alacağız (diğer durumlar ile ilgili tartışma için kitaba bakınız). $H_0 : \mu_X = \mu_Z$ altında

$$T = \frac{\bar{X} - \bar{Z}}{\sqrt{\frac{\sigma_X^2}{n_1} + \frac{\sigma_Z^2}{n_2}}} \sim N(0, 1)$$

Sezgisel olarak, eğer boş hipotez doğru değil ise, T büyük olmalı (mutlak değer cinsinden). Dolayısıyla, α büyüklüğe sahip H_A 'ya karşı H_0 testinde H_0 ret edilir eğer aşağıdaki koşul sağlanırsa:

$$|T| > -\Phi^{-1}\left(\frac{\alpha}{2}\right)$$

2. Varyans testi için, dağılımlar ile ilgili sonuçları hatırlamak gerekir:

$$\frac{(n_1 - 1)s_X^2}{\sigma_X^2} \sim \chi_{n_1 - 1}^2$$

ve

$$\frac{(n_2 - 1)s_Z^2}{\sigma_Z^2} \sim \chi_{n_2 - 1}^2$$

Bunlar birbirinden bağımsızdır. Hatırlayınız, bağımsız ki-karelerin serbestlik derecesi ile bölümü F dağılımlıydı:

$$S' = \frac{\frac{(n_1 - 1)s_X^2}{\sigma_X^2} / (n_1 - 1)}{\frac{(n_2 - 1)s_Z^2}{\sigma_Z^2} / (n_2 - 1)} \sim F_{n_1 - 1, n_2 - 1}$$

Açıkçası biz σ_X^2 ile σ_Z^2 'yi bilmiyoruz, fakat $H_0 : \sigma_X^2 = \sigma_Z^2$ altında bu ifade basitleştirilebilir,

$$\bar{S} = \frac{s_X^2}{s_Z^2}$$

Dolayısıyla, eğer aşağıdaki koşullar sağlanırsa α büyüklüklü bir test ret eder:

$$\begin{aligned} \tilde{S} &> F_{n_1-1, n_2-1}(1 - \alpha/2) \quad \text{veya} \\ \tilde{S} &< F_{n_1-1, n_2-1}(\alpha/2) \end{aligned}$$

2.2 Parametrik Olmayan Çıkarımlar

Şimdiye kadar, veri üretim prosesi $f(x|\theta)$ formunda olan ve sonlu parametre θ boyutuna kadar bilinen durumdaki problemler ile ilgilendik. O durumdaki testlere *parametrik çıkarım* denilir.

İstisna olarak, tahmin konusunda vurguladığımız gibi, örneklem ortalamaları, varyansları ve diğer momentleri *herhangi* bir dağılımın ortalamalarının, varyanslarının ve diğer yüksek-sıralı momentlerinin hesaplanması için avantajlı özelliklere sahiptir.

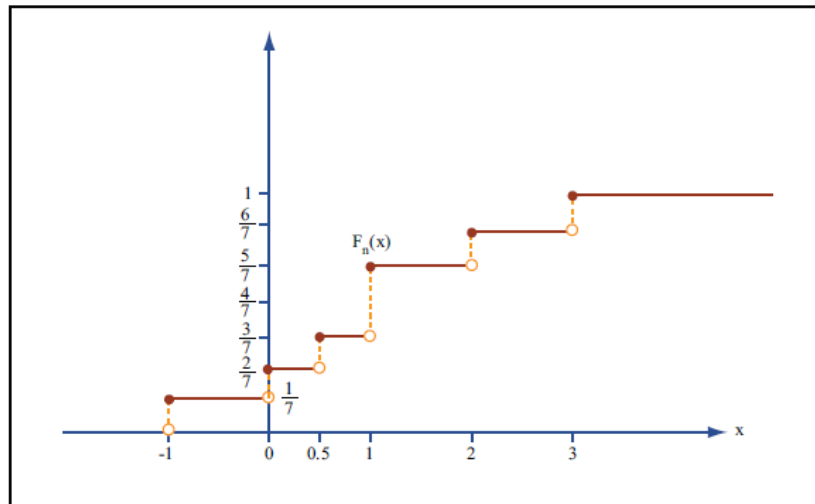
Bir rasgele değişkenin dağılımının tümü kendi c.d.f.si ile karakterize edilebildiği için, herhangi bir sınırlama getirmeden(elbette ki geçerli bir c.d.f. olmalı yani monoton ve sağdan sürekli), sanki veriden c.d.f.yi tahmin etmek iyi bir fikirmiş gibi görünüyor.

Örnek dağılım fonksiyonu $F_n(x)$ şöyledir:

$$F_n(x) = \frac{j}{n} \text{ for } X_{(j)} \leq x < X_{(j+1)}$$

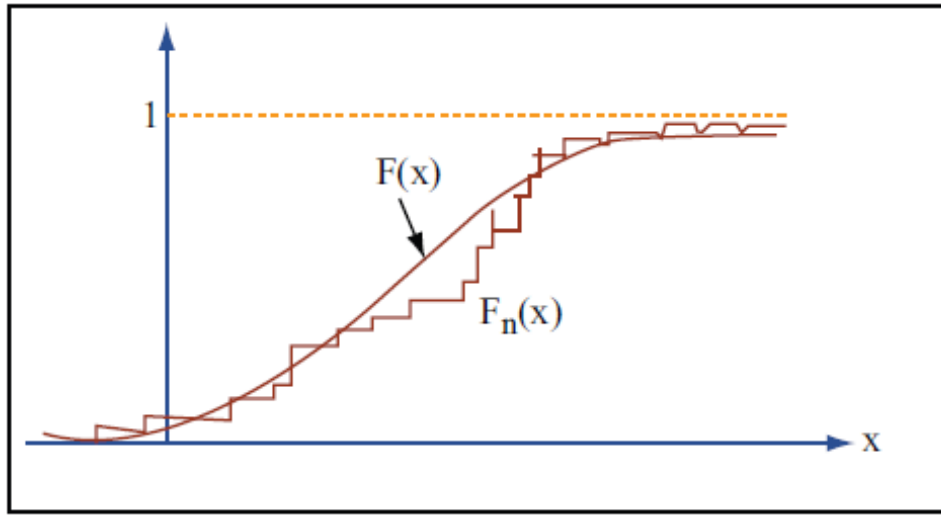
Burada $X_{(j)}$ j. sıralı istatistiktir(bunun örneklemdeki j en küçük değer olduğunu hatırlayınız), ayrıca $X_{(0)} \equiv -\infty$ ve $X_{(n+1)} \equiv \infty$ 'dur.

Örnek 6. bir $\{-1, 3, 1, 1, 0.5, 2, 0\}$ örneklemini için, sıralanmış örneklem $\{-1, 0, 0.5, 1, 1, 2, 3\}$ 'tür ve örneklem dağılım fonksiyonu $F_n(x)$ 'i grafik ile gösterebiliriz:



Kaynak: MIT OpenCourseWare

Bilinmeyen bir dağılım ailesinden elde edilen bir X_1, \dots, X_n rasgele örneklem ile ilgili çıkarım problemi ile ilgileniyoruz ve c.d.f.si $F(x)$ olan (örneğin bir standart normal dağılım için $F(x) = \Phi(x)$ gibi) belirli bir dağılımdan elde edilip edilmediğini test etmeyi arzuluyoruz. Daha önceki tartışmalarda altı çizilen testlerden herhangi birini uygulamak için elimizde spesifik parametreler olmadığı için, test fikri $F_n(X)$ 'in $F(x)$ 'ten “çok fazla” sapıp saptığını kontrol etmek olur.



Kaynak: MIT OpenCourseWare

2.3. Kolmogorov-Smirnov Testi

Gözlemlenen bir örneklem $F(x)$ dağılımı tarafından türetilip türetilmediğin test etmek için, Kolmogorov-Smirnov istatistiğinin büyük değerleri için testi ret ederiz. İstatistik aşağıdaki gibi tanımlanır,

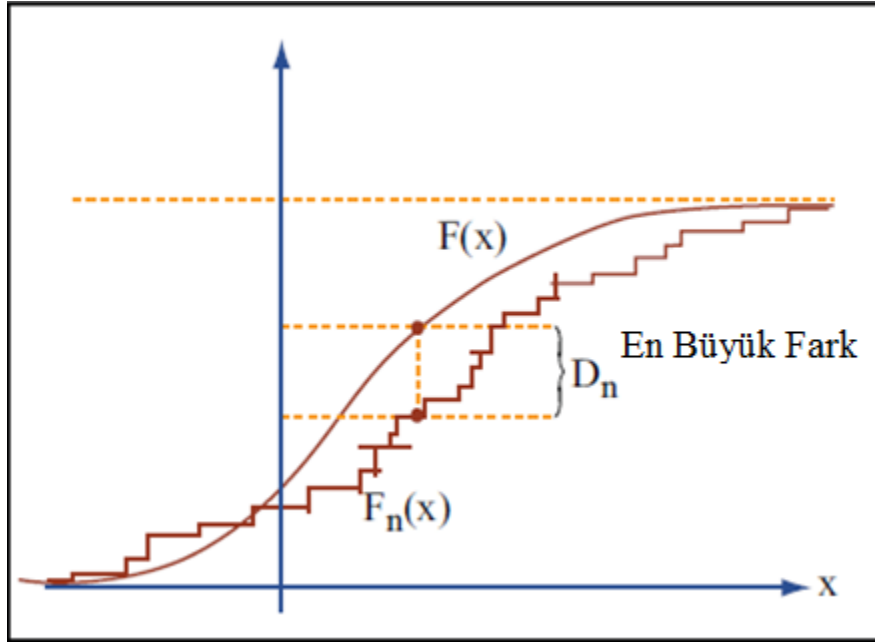
$$D_n = \sup_x |F_n(x) - F(x)|$$

burada $\sup_x F(x)$ supremumdur, yani $\{F(x): x \in \mathbb{R}\}$ 'in en küçük üst sınırıdır – ufak kümelerde sürekli fonksiyonlar için bu maksimumun aynısıdır, fakat Kolmogorov-Smirnov istatistiği $1/n$ büyüklüğündeki sıçramaları içeren örneklem dağılım fonksiyonu içerdiği ve supremum bütün reel sayı çizgisini kapsadığı için, gerçekte hiçbir belirli x değerine ulaşılmamış olabilir.

İstatistiğin kritik değerleri asimptotik (yani büyük n'ler için) dağılım fonksiyonundan elde edilebilir

$$G(D_n) = \lim_{n \rightarrow \infty} P(D_n < x) = 1 - 2 \sum_{i=1}^{\infty} (-1)^{i-1} e^{-2i^2 x^2}$$

Bu ifadeyi doğuran argümanlar çok açık değildir ve çok tekniktir çünkü bu, reel sayıları değil, fonksiyonların dağılımını içeriyor (rasgele fonksiyonlar genellikle *stokastik proses* olarak adlandırılır).



Kaynak: MIT OpenCourseWare

C.d.f.nin hesaplanması sonsuz bir serinin hesaplanmasını gerektirdiği için, formül kullanmak basit değildir. Ancak, birçok ders kitabı yaygın kritik değerler için tablo oluşturur.

Örnek 7. Varsayalım ki bir madeni parayı tekrar tekrar, örneğin 160 kere, fırlatıyoruz ve örneklemin bir $B(4, 0.5)$ dağılımından türetilip türetilmediğin $\alpha = 0.2$ güvenirlilik düzeyi ile test etmek istiyoruz. Diyelim ki aşağıdaki örneklem frekanslarını gözlemledik:

Tura sayısı	0	1	2	3	4
Örneklem frekansı	10	33	61	43	13
Kümülatif örneklem frekansı $F_n(\cdot)$	10	43	104	147	160
H_0 $F(\cdot)$ altında Kümülatif frekans fark	10	50	110	150	160
	0	7	6	3	0

O zaman Kolmogorov-Smirnov istatistiği şuna eşittir:

$$D_n = \frac{1}{160} \max\{0.7, 6, 3, 0\} = \frac{7}{160} \approx 0.044$$

Kitaptaki asimptotik formülünü kullanarak, $C_{0.20} = \frac{1.07}{\sqrt{160}} \approx 0.85$ 'tir. $0.44 < 0.85$ olduğu için, boş hipotezi %20 düzeyinde ret edemiyoruz.

2.4. 2-Örneklemlili Kolmogorov-Smirnov Testi

Varsayalım ki dağılım ailesi bilinmeyen iki bağımsız rasgele örneklemimiz var ve bunlar X_1, \dots, X_n ile Y_1, \dots, Y_n olsun. Her iki örnekleminde aynı dağılım tarafından türetilip türetilmediğini test etmek istiyoruz. Buradaki düşünce $F_n(x)$ ile $G_n(x)$ 'in birbirinden “çok uzakta” olup olmadığını test etmektir.

Bir istatistik oluşturuyoruz,

$$D = \sup_x |F_n(x) - G_n(x)|$$

ve büyük D değerleri için ret ediyoruz. Eğer aşağıdaki koşul sağlanırsa, α büyüklüğündeki bir test için kritik değerlerin asimptotik olarak iyi bir tahmini testi ret eder:

$$D > \sqrt{-\frac{1}{2} \left(\frac{1}{m} + \frac{1}{n} \right) \log \frac{\alpha}{2}}$$

2.5 Pearson'nın χ^2 Testi

Varsayalım ki n tane i.i.d. gözlemlili bir örneklemden her bir X_i değeri k kadar, A_1, \dots, A_k , kategoriden birine yerleşecek şekilde sınıflandırıldı. p_1, \dots, p_k her bir kategorinin olasılığı, ve f_1, \dots, f_k 'de gözlemlenen frekanslar olsun. Aşağıdaki bileşik hipotezi

$$H_0 : p_1 = \pi_1, p_2 = \pi_2, \dots, p_k = \pi_k$$

alternatife karşı test ettiğimizi düşünelim. Alternatif hipotez bu eşitliklerden en az iki veya daha fazlasının gerçekleşmemesidir (olasılıkların toplamı 1 olduğu için, burada tam olarak bir eşitlik tutmadı diyemeyiz). Aşağıdaki istatistiği kullanabiliriz

$$T = \sum_{i=1}^k \frac{(f_i - n\pi_i)^2}{n\pi_i}$$

ve büyük T deęerleri için ret ederiz. En uygun kritik deęerleri belirlemek için, T 'nin nasıl dağıldığını bilmek zorundayız. Maalesef, bu dağılım ilgili modele baęlıdır. Ancak, H_0 altında dağılım asimptotik olarak modelden bağımsızdır ve büyük n örneklem için asimptotik olarak $T \sim \chi^2_{k-1}$ 'dir. Pratik bir kural olarak, eęer $n \geq 4k$ ise ki-kare tahmini çalışabilir.