ÖRNEKLEM DAĞILIMLARI (Kİ-KARE, F, T)

Doç. Dr. Yasemin Kayhan Atılgan (Şube 01)
Doç. Dr. Derya Ersel (Şube 02)

Parametresi θ ve olasılık / olasılık yoğunluk fonksiyonu $f(x;\theta)$ olan bir kitleden çekilen rasgele örneklem $X_1, X_2, \ldots, X_n$ olsun. $T = T(X_1, X_2, \ldots, X_n)$, bu rasgele örneklem üzerinden tanımlanmış bir istatistik olsun. İstatistik, gözlenen raslantı değişkenlerinin ve bilinen sabitlerin bir fonksiyonudur. Kitle dağılımı biliniyorsa T istatistiğinin dağılımı kolaylıkla bulunabilir. Daha önce bu istatistiklerin dağılımını bulabilmek için kullanılan çeşitli yöntemler üzerinde durulmuştu.

T istatistiğinin olasılık dağılımına "T'nin örneklem dağılımı" denir. Bu örneklem dağılımları bilinmeyen kitle parametrelerini tahmin etmek ve bu parametreler ile karar vermek amacıyla kullanılabilir.

Örneğin, μ kitle ortalaması tahmin edilmek istensin. Bu kitleden çekilen $X_1, X_2, \ldots, X_n$ örnekleminden yararlanarak bu kitle ortalaması,

$$\bar{X} = \sum_{i=1}^{n} \frac{X_i}{n}$$

örneklem ortalaması ile tahmin edilebilir. Dikkat edilmelidir ki buradaki $\bar{X}$ raslantı değişkeni $X_1, X_2, \ldots, X_n$ raslantı değişkenlerinin bir fonksiyonudur. Dolayısıyla bir istatistiktir. Örneklem ortalamasının dışında S^2 örneklem varyansı, $Y_n = \max(X_1, X_2, \ldots, X_n)$, $Y_1 = \min(X_1, X_2, \ldots, X_n)$, $R = Y_n - Y_1$, örneklem ortancası da birer istatistiktir. Bu istatistikler aynı zamanda birer raslantı değişkeni oldukları için olasılık dağılımları elde edilebilir. Bu olasılık dağılımları, örneklem dağılımlarıdır.

Doğrusal birleşim, Chebyshev Eşitsizliği ve Merkezi Limit Teoremi konularında özellikle örneklem ortalamasının örneklem dağılımını bulmada kullanılan bir çok yöntem üzerinde duruldu.

Normal dağılım, istatistikte önemli bir dağılım olduğu için, normal dağılımlı kitlelerden alınan örneklemler üzerinden tanımlı istatistiklerin örneklem dağılımları çok önemlidir. Normal dağılımdan alınan örneklemler için en çok kullanılan örneklem dağılımları, ki-kare, t ve F dağılımıdır.

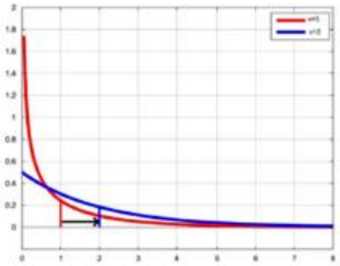
Kİ-KARE DAĞILIMI

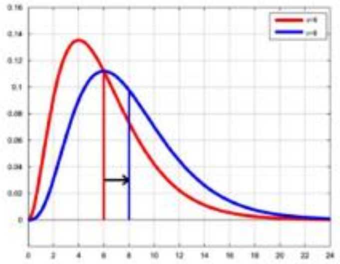
TANIM: $\nu \in \mathbb{N}$ olmak üzere, destek kümesi $R_X = [0, \infty)$ biçiminde pozitif gerçel sayılar olan bir X raslantı değişkeninin olasılık yoğunluk fonksiyonu aşağıdaki gibi yazılabiliyorsa bu raslantı değişkeni ν serbestlik derecesi ile ki-kare dağılımına sahiptir $\left(X \sim \chi^2_\nu\right)$.

$$f(x) = \begin{cases} \dfrac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\nu/2-1} e^{-x/2}, & x > 0 \text{ için} \\ 0, & \text{ö.d. için} \end{cases}$$

Farklı serbestlik derecelerine sahip ki-kare dağılımının biçimleri aşağıdaki grafiklerde verilmiştir.

Grafiklerdeki dikey çizgiler o dağılımın ortalamasını göstermektedir.





- $X \sim \chi^2_\nu$ için $E(X) = \nu$ biçimindedir. Ki-kare dağılımının beklenen değeri serbestlik derecesine eşittir.

İspat: $E(X) = \int_0^\infty x f(x) dx = \int_0^\infty x \dfrac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\nu/2-1} e^{-x/2} dx = \dfrac{1}{2^{\nu/2}\Gamma(\nu/2)} \int_0^\infty x^{\nu/2} e^{-x/2} dx$

$\dfrac{x}{2} = u$ dönüşümü yapılsın. Bu durumda, $dx = 2du$ olacaktır.

$$E(X) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} \int_0^\infty (2u)^{\nu/2} e^{-u} 2du = \frac{2}{\Gamma(\nu/2)} \int_0^\infty u^{\nu/2} e^{-u} du$$

Gamma integrali: $\Gamma(n) = \int_0^\infty x^{n-1} e^{-x} dx$ olduğundan,

$$E(X) = \frac{2}{\Gamma(\nu/2)} \underbrace{\int_0^\infty u^{\nu/2} e^{-u} du}_{\Gamma\left(\frac{\nu}{2}+1\right)} = \frac{2}{\Gamma\left(\frac{\nu}{2}\right)} \Gamma\left(\frac{\nu}{2}+1\right) = \frac{2\left(\frac{\nu}{2}\right)\Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} = \nu$$

- $X \sim \chi^2_\nu$ için $V(X) = 2\nu$ biçimindedir.

İspat: $V(X) = E(X^2) - [E(X)]^2$

$$E(X^2) = \int_0^\infty x^2 f(x) dx = \int_0^\infty x^2 \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\nu/2-1} e^{-x/2} dx = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} \int_0^\infty x^{\nu/2+1} e^{-x/2} dx$$

$\dfrac{x}{2} = u$ dönüşümü yapılsın. Bu durumda, $dx = 2du$ olacaktır.

$$E(X^2) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} \int_0^\infty (2u)^{\nu/2+1} e^{-u} 2du = \frac{4}{\Gamma(\nu/2)} \int_0^\infty u^{\nu/2+1} e^{-u} du$$

$$E(X^2) = \frac{4}{\Gamma\left(\frac{\nu}{2}\right)} \Gamma\left(\frac{\nu}{2}+2\right) = \frac{4\left(\frac{\nu}{2}+1\right)\left(\frac{\nu}{2}\right)\Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} = \nu(\nu+2)$$

$V(X) = \nu(\nu+2) - \nu^2 = 2\nu$

- $X \sim \chi^2_\nu$ için $M_X(t) = (1-2t)^{-\nu/2}$'dir.

Bu moment çıkaran fonksiyonun tanımlı olabilmesi için $\frac{1}{2} - t > 0$ yani $t < \frac{1}{2}$ olmalıdır.

Ki-kare dağılımı ile Gamma dağılımı arasındaki ilişki:

Ki-kare dağılımı ile Gamma dağılımı arasındaki ilişkiyi vermeden önce Gamma dağılımı ve Gamma fonksiyonu üzerinde kısaca duralım.

TANIM: Sürekli bir X raslantı değişkeninin olasılık yoğunluk fonksiyonu aşağıdaki gibi yazılabiliyorsa bu raslantı değişkeni $\alpha > 0$ ve $\beta > 0$ için Gamma dağılımına sahiptir $(X \sim Gamma(\alpha, \beta))$.

$$f(x) = \begin{cases} \dfrac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}, & x > 0 \text{ için} \\ 0, & \text{ö.d. için} \end{cases}$$

Farklı α ve β parametreleri için Gamma dağılımının biçimi aşağıdaki şekilde verilmiştir.

TANIM: Gamma fonksiyonu, faktöriyel işleminin tamsayı olmayan sayılar için genelleştirilmiş halidir ve aşağıdaki gibi tanımlanır.

$$\Gamma(n) = \int_0^\infty x^{n-1} e^{-x} dx$$

Gamma fonksiyonunun bazı özellikleri aşağıdaki gibi verilebilir.

1) $\Gamma(1) = 1$'dir.
2) $n > 1$ için $\Gamma(n) = (n-1)\Gamma(n-1)$ 'dir.
3) $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ 'dir.
4) $\Gamma\left(-\frac{1}{2}\right) = -2\sqrt{\pi}$ 'dir.
5) $n \in \mathbb{N}$ ise $\Gamma(n) = (n-1)!$'dir.

ÖRNEK: $\Gamma\left(\frac{5}{2}\right) = \left(\frac{5}{2}-1\right)\Gamma\left(\frac{5}{2}-1\right) = \left(\frac{3}{2}\right)\Gamma\left(\frac{3}{2}\right) = \left(\frac{3}{2}\right)\left(\frac{3}{2}-1\right)\Gamma\left(\frac{3}{2}-1\right)$

$$= \left(\frac{3}{2}\right)\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right) = \frac{3}{4}\sqrt{\pi}$$

TEOREM : α ve β parametreleri ile Gamma dağılımına sahip bir X raslantı değişkeninin beklenen değer, varyans ve moment çıkaran fonksiyonu aşağıdaki gibidir.

$$E(X) = \alpha\beta \qquad V(X) = \alpha\beta^2 \qquad M_X(t) = (1-\beta t)^{-\alpha}$$

NOT: $\alpha = v/2$ ve $\beta = 2$ için Gamma dağılımı, ki-kare dağılımına dönüşür. Bu durumda, ki-kare dağılımının Gamma dağılımının özel bir durumu olduğu söylenebilir.

$X \sim Gamma(\alpha, \beta)$ için $M_X(t) = (1-\beta t)^{-\alpha}$ olduğundan, $\alpha = v/2$ ve $\beta = 2$ ise $M_X(t) = (1-2t)^{-\frac{v}{2}}$ olur.

Böylece v serbestlik dereceli ki-kare dağılımının moment çıkaran fonksiyonuna ulaşılmış olur. Benzer şekilde, $\alpha = v/2$ ve $\beta = 2$ için

$$E(X) = \alpha\beta = \frac{v}{2}\cdot 2 = v \quad \text{ve} \quad V(X) = \alpha\beta^2 = \frac{v}{2}\cdot 2^2 = 2v$$

elde edilir. Böylece v serbestlik dereceli ki-kare dağılımının beklenen değer ve varyansına ulaşılmış olur.

Ki-kare Dağılımı ile İlgili Teoremler:

TEOREM 1: X raslantı değişkeni ortalaması μ, varyansı σ^2 olan normal dağılıma sahip olmak üzere,

$$Z^2 = \left(\frac{X-\mu}{\sigma}\right)^2$$

biçiminde tanımlanan raslantı değişkeni 1 serbestlik derecesi ile ki-kare dağılımına sahip olur $\left(Z^2 \sim \chi_1^2\right)$.

İspat: $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \quad -\infty < x < \infty$

$y = x^2$ dönüşümü yapılsın. Bu durumda, $x = \pm\sqrt{y}, \quad 0 < y < \infty$

$$g(y) = \frac{1}{2\sqrt{y}}\left[f_X\left(\sqrt{y}\right) + f_X\left(-\sqrt{y}\right)\right] = \frac{1}{2\sqrt{y}}\left[\frac{1}{\sqrt{2\pi}}e^{-\frac{y}{2}} + \frac{1}{\sqrt{2\pi}}e^{-\frac{y}{2}}\right]$$

$$g(y) = \frac{1}{\Gamma\left(\frac{1}{2}\right)2^{1/2}} y^{\frac{1}{2}-1} e^{-y/2}, \; y > 0 \;\Rightarrow\; Y \sim \chi_1^2$$

TEOREM 2: $Z_1, Z_2, \ldots, Z_n$ standart normal dağılıma sahip bağımsız raslantı değişkenleri olmak üzere,

$$\sum_{i=1}^n Z_i^2 = Z_1^2 + Z_2^2 + \cdots + Z_n^2$$

biçiminde tanımlanan raslantı değişkeni n serbestlik derecesi ile ki-kare $\left(\chi_n^2\right)$ dağılımına sahip olur.

İspat: 1. Teorem'den her $i = 1, 2, \ldots, n$ için $Z_i^2 \sim \chi_1^2$ olduğu biliniyor. O halde, her bir Z_i^2 değişkeninin moment çıkaran fonksiyonu,

$$M_{Z_i^2}(t) = (1-2t)^{-1/2}, \qquad i = 1, 2, \ldots, n$$

biçimindedir. MÇF yönteminden,

$$M_{\sum_{i=1}^n Z_i^2}(t) = \prod_{i=1}^n M_{Z_i^2}(t) = \prod_{i=1}^n (1-2t)^{-1/2} = (1-2t)^{-n/2}$$

bulunur. Bu da n s.d.'li ki-kare $\left(\chi_n^2\right)$ dağılımının MÇF'udur.

TEOREM 3: $X_1, X_2, \ldots, X_n$ raslantı değişkenleri, sırasıyla $v_1, v_2, \ldots, v_n$ serbestlik dereceleri ile ki-kare dağılımına sahip bağımsız raslantı değişkenleri olmak üzere,

$$Y=\sum_{i=1}^{n} X_i$$

biçiminde tanımlanan raslantı değişkeni $v=\sum_{i=1}^{n} v_i$ serbestlik derecesi ile ki-kare dağılımına sahip olur.

İspat: Her bir $X_i \sim \chi^2_{v_i}$, $i=1,2,\ldots,n$ değişkeninin moment çıkaran fonksiyonu,

$$M_{X_i}(t)=(1-2t)^{-v_i/2}$$

olarak yazılabilir. MÇF yönteminden,

$$M_Y(t)=M_{\sum X_i}(t)=\prod_{i=1}^{n} M_{X_i}(t)=\prod_{i=1}^{n}(1-2t)^{-v_i/2}=(1-2t)^{-\sum_{i=1}^{n} v_i/2}$$

bulunur. Bu da $v=\sum_{i=1}^{n} v_i$ serbestlik dereceli ki-kare dağılımının MÇF'udur.

TEOREM 4: X_1 raslantı değişkeni v_1, X_1+X_2 raslantı değişkeni v serbestlik derecesi ile ki-kare dağılımına sahip raslantı değişkenleri olmak üzere, X_2 raslantı değişkeni $v-v_1$ serbestlik derecesi ile ki-kare dağılımına sahip olur.

İspat: $X_1 \sim \chi^2_{v_1}$ için $M_{X_1}(t)=(1-2t)^{-v_1/2}$

$X_1+X_2 \sim \chi^2_{v}$ için $M_X(t)=(1-2t)^{-v/2}$

MÇF yönteminden,

$$M_X(t)=M_{X_1}(t)M_{X_2}(t)$$

$$(1-2t)^{-v/2}=(1-2t)^{-v_1/2}M_{X_2}(t) \Rightarrow M_{X_2}(t)=(1-2t)^{-(v-v_1)/2}$$

bulunur. Böylece X_2 raslantı değişkeninin $v-v_1$ serbestlik derecesi ile ki-kare dağılımına sahip olduğu söylenebilir.

TEOREM 5: $X_1, X_2, \ldots, X_n$, $N(\mu,\sigma^2)$ normal dağılımdan alınan n büyüklükte rasgele örneklemin ortalaması $\bar{X}=\sum_{i=1}^{n}\frac{X_i}{n}$, varyansı $S^2=\sum_{i=1}^{n}\frac{(X_i-\bar{X})^2}{n-1}$ olsun.

a. $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

b. $\bar{X}$ ve S^2 bağımsızdır.

İspat: Teroem 5-(a)'yı n=2 durumu için ispatlayalım.

$X_1 \sim N(\mu,\sigma^2)$ ve $X_2 \sim N(\mu,\sigma^2)$ olduğundan,

$X_1+X_2 \sim N(2\mu,2\sigma^2)$ ve $X_1-X_2 \sim N(0,2\sigma^2)$ olur.

X_1-X_2 raslantı değişkeni standartlaştırılırsa, $\frac{X_1-X_2}{\sqrt{2\sigma^2}} \sim N(0,1)$ elde edilir.

SND'a sahip bu raslantı değişkeninin karesi 1 serbestlik derecesi ile ki-kare dağılımına sahip olur.

$$\left(\frac{X_1-X_2}{\sqrt{2\sigma^2}}\right)^2=\frac{1}{2}\frac{(X_1-X_2)^2}{\sigma^2} \sim \chi^2_1$$

$n=2$ için $\frac{(n-1)S^2}{\sigma^2}$ ifadesini açalım. Bunun için ilk olarak n=2 için S^2 terimini yazalım.

$$S^2=\sum_{i=1}^{2}\frac{(X_i-\bar{X})^2}{2-1}=(X_1-\bar{X})^2+(X_2-\bar{X})^2=\left(X_1-\frac{X_1+X_2}{2}\right)^2+\left(X_2-\frac{X_1+X_2}{2}\right)^2$$

$$=\left(\frac{X_1}{2}-\frac{X_2}{2}\right)^2+\left(\frac{X_2}{2}-\frac{X_1}{2}\right)^2=\left[\frac{1}{2}(X_1-X_2)\right]^2+\left[\frac{1}{2}(X_2-X_1)\right]^2$$

$$=2\left[\frac{1}{2}(X_1-X_2)\right]^2=\frac{(X_1-X_2)^2}{2}$$

$$\frac{(n-1)S^2}{\sigma^2}=\frac{(2-1)}{\sigma^2}\frac{(X_1-X_2)^2}{2}=\frac{1}{2}\frac{(X_1-X_2)^2}{\sigma^2}$$

Bu ifade, yukarıda dağılımını χ^2_1 olarak bulduğumuz raslantı değişkeni ile aynıdır. Böylece $n=2$ için Teroem 5-(a) ispatlanmış olur.

$n=2$ için Teorem 5-(b)'yi ispatlayalım.

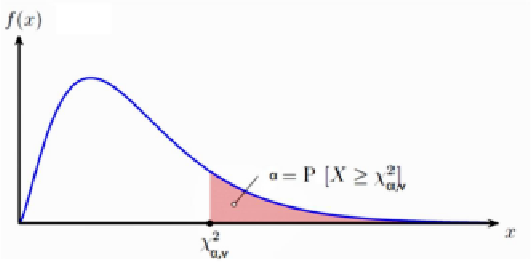
X_1 ve X_2 bağımsız olduğundan ,

$$Cov(X_1+X_2, X_1-X_2) = Cov(X_1,X_1) - Cov(X_1,X_2) + Cov(X_2,X_1) - Cov(X_2,X_2)$$
$$= \sigma^2 - 0 + 0 - \sigma^2 = 0$$

elde edilir. Böylece X_1+X_2 ve X_1-X_2 iki değişkenli normal dağılıma sahip ilişkisiz raslantı değişkenleridir. Sonuç olarak, bu raslantı değişkenleri bağımsızdır.

X_1+X_2 ve X_1-X_2 bağımsızsa, $\frac{1}{2}(X_1+X_2)$ ile $\frac{(X_1-X_2)^2}{2}$ raslantı değişkenleri de bağımsızdır. Böylece, $\bar{X}$ ve S^2'nin bağımsız olduğu gösterilmiş olur.

Ki-kare Tablosu



Ki-kare tablosu, yukarıdaki şekilden de anlaşılabileceği gibi, sağ tarafındaki alan α olan $\chi^2_{\alpha,v}$ değerini verir. Tablonun sol tarafında v serbestlik dereceleri, tablonun üstünde ise α olasılıkları yer alır.

$X \sim \chi^2_v$ olsun. Bu durumda ki-kare tablosu α olasılığını ya da $\chi^2_{\alpha,v}$ değerini aşağıdaki eşitlikten yararlanarak bulmamızı sağlar.

$$\alpha = P(X > \chi^2_{\alpha,v}) = \int_{\chi^2_{\alpha,v}}^{\infty} f(x)dx$$

ÖRNEK: X_1 ve X_2 aynı $N(0,100)$ normal dağılımına sahip bağımsız raslantı değişkenleri olmak üzere $P(X_1^2+X_2^2>10)$ olasılığını hesaplayınız.

Çözüm: $X_1 \sim N(0,100)$ ve $X_2 \sim N(0,100)$ raslantı değişkenlerini standartlaştırırsak.

$Z_1 = \frac{X_1-0}{10} = \frac{X_1}{10} \sim N(0,1)$ ve $Z_2 = \frac{X_2-0}{10} = \frac{X_2}{10} \sim N(0,1)$ elde edilir.

Ki-kare Teorem 1'den, $Z_1^2 = \frac{X_1^2}{100} \sim \chi_1^2$ ve $Z_2^2 = \frac{X_2^2}{100} \sim \chi_1^2$ elde edilir.

Ki-kare Teorem 2'den, $Z_1^2+Z_2^2 = \frac{X_1^2}{100} + \frac{X_2^2}{100} = \frac{X_1^2+X_2^2}{100} \sim \chi_2^2$ elde edilir.

Ki-kare tablosundan,

$$P(X_1^2+X_2^2>10) = P\left(\frac{X_1^2+X_2^2}{100} > \frac{10}{100}\right) = P(\chi_2^2 > 0{,}1) = 0{,}95$$

ÖRNEK: X_1, X_2, X_3, X_4, aynı $N(\mu,9)$ normal dağılımdan çekilen rasgele örneklem olsun. $P(S^2 \le k) = 0{,}05$ olmak üzere k değerini bulunuz.

Çözüm: Ki-kare Teorem 5'ten,

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}, \quad n=4, \quad \sigma^2 = 9 \Rightarrow \frac{3S^2}{9} \sim \chi_3^2$$

$$P(S^2 \le k) = 0{,}05 \Rightarrow P\left(\frac{3S^2}{9} \le \frac{3k}{9}\right) = 0{,}05 \Rightarrow P\left(\chi_3^2 \le \frac{3k}{9}\right) = 0{,}05$$

$$\Rightarrow P\left(\chi_3^2 > \frac{3k}{9}\right) = 0{,}95 \Rightarrow \frac{3k}{9} = 0{,}352 \Rightarrow k = 1{,}056$$

ÖRNEK: $X_1, X_2, \dots, X_n$, aynı $N(\mu, \sigma^2)$ normal dağılımlı bağımsız raslantı değişkenleri ve $Z \sim N(0,1)$ raslantı değişkeni verilsin. $X_1, X_2, \dots, X_n$ r.d.'leri ile Z r.d. bağımsız olsun. $W = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2 + Z^2$ biçiminde tanımlanan raslantı değişkeninin dağılımını bulunuz.

Çözüm:

$X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$ ise,

$Z_i = \frac{X_i - \mu}{\sigma} \sim N(0,1), \quad i = 1,2,\dots,n$ olur.

Ki-kare Teorem 1'den, $Z_i^2 = \left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi_1^2, \quad i=1,2,\dots,n$ ve $Z^2 \sim \chi_1^2$,

Ki-kare Teorem 2'den, $\sum_{i=1}^{n} Z_i^2 = \sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2} \sim \chi_n^2$ olur.

Ki-kare Teorem 3'ten yararlanarak,

$$W = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2 + Z^2 \sim \chi_{n+1}^2$$

ÖRNEK: $N(\mu, \sigma^2)$ dağılımından n birimlik rasgele örneklemin varyansı

$S^2 = \sum_{i=1}^{n}\frac{(X_i-\bar{X})^2}{n-1}$ olmak üzere, $E(S^2)$ ve $V(S^2)$'yi hesaplayınız.

Çözüm: Ki-kare Teorem 5'ten,

$Y = \frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$ olduğundan, $E(Y) = n-1$, $V(Y) = 2(n-1)$ yazılabilir.

Buradan,

$$E(Y) = n-1 \Rightarrow E\left[\frac{(n-1)S^2}{\sigma^2}\right] = n-1 \Rightarrow \frac{n-1}{\sigma^2}E(S^2) = n-1 \Rightarrow E(S^2) = \sigma^2$$

$$V(Y) = 2(n-1) \Rightarrow V\left[\frac{(n-1)S^2}{\sigma^2}\right] = 2(n-1) \Rightarrow \frac{(n-1)^2 S^2}{\sigma^4}V(S^2) = 2(n-1) \Rightarrow V(S^2) = \frac{2\sigma^{42}}{n-1}$$

bulunur.

F DAĞILIMI

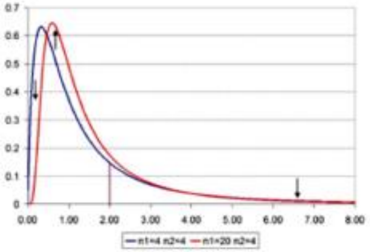
F dağılımı, istatistiğin birçok uygulamasında kullanılır. Özellikle, varyans analizinde kitle varyanslarının eşitliğinin test edilmesine ilişkin geliştirilen yöntemler bu dağılıma dayanır.

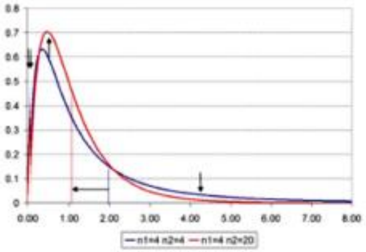
TANIM: Sürekli X raslantı değişkeninin olasılık yoğunluk fonksiyonu aşağıdaki biçimde yazılabiliyorsa bu raslantı değişkeni v_1 ve v_2 serbestlik dereceleri ile F dağılımına sahiptir $(X \sim F_{v_1,v_2})$.

$$f(x) = \begin{cases} \dfrac{\Gamma\left(\frac{v_1+v_2}{2}\right)\left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2}} x^{\frac{v_1}{2}-1}}{\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)\left(1+\frac{v_1}{v_2}x\right)^{\frac{v_1+v_2}{2}}}, & x > 0 \text{ için} \\ 0, & \text{ö.d. için} \end{cases}$$

X raslantı değişkeninin v_1 ve v_2 serbestlik dereceleri ile F dağılımına sahip olduğunu göstermek için $X \sim F_{v_1,v_2}$ ifadesi kullanılır.

F dağılımının farklı serbestlik derecelerine göre grafiği aşağıda verilmiştir. Grafikteki dikey çizgiler dağılımın ortalamasını göstermektedir.





Beta Dağılımı ile F Dağılımı Arasındaki İlişki

F dağılımı Beta dağılımından elde edilir.

TANIM: Destek kümesi $R_X = [0,1]$ olan sürekli bir X raslantı değişkeninin olasılık yoğunluk fonksiyonu $\alpha, \beta \in R^{+}$ için aşağıdaki gibi yazılıyorsa bu raslantı değişkeni Beta dağılımına sahiptir $(X \sim Beta(\alpha, \beta))$.

$$f(x) = \begin{cases} \dfrac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}, & 0 < x < 1 \text{ için} \\ 0, & \text{ö.d. için} \end{cases}$$

Beta fonksiyonu, $B(\alpha,\beta) = \dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ olarak tanımlanır. Beta fonksiyonunun farklı integral gösterimleri de vardır. Bunlar,

1. $B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx$
2. $B(\alpha,\beta) = \int_0^\infty x^{\alpha-1}(1+x)^{-\alpha-\beta}\,dx$

NOT: Beta dağılımında $\alpha = \frac{\nu_1}{2}$ ve $\beta = \frac{\nu_2}{2}$ alınır ve $y = \dfrac{\nu_2 x}{\nu_1(1-x)}$ dönüşümü yapılırsa Y raslantı değişkeninin dağılımı F_{ν_1,ν_2} olur.

$Beta\left(\frac{\nu_1}{2}, \frac{\nu_2}{2}\right)$ dağılımı: $f(x) = \dfrac{1}{B\left(\frac{\nu_1}{2}, \frac{\nu_2}{2}\right)} x^{\frac{\nu_1}{2}-1}(1-x)^{\frac{\nu_2}{2}-1}, \quad 0 < x < 1$

Dönüşüm yönteminden,

$$f(y) = f(x)\left|\frac{dx}{dy}\right|$$

$y = \dfrac{\nu_2 x}{\nu_1(1-x)} \Rightarrow x = \dfrac{\nu_1 y}{\nu_2 + \nu_1 y}$ olur. Buradan, $\dfrac{dx}{dy} = \dfrac{\nu_1 \nu_2}{(\nu_2 + \nu_1 y)^2}$ elde edilir.

$$f(y) = f(x)\left|\frac{dx}{dy}\right| = \frac{1}{B\left(\frac{\nu_1}{2}, \frac{\nu_2}{2}\right)}\left(\frac{\nu_1 y}{\nu_2 + \nu_1 y}\right)^{\frac{\nu_1}{2}-1}\left(1 - \frac{\nu_1 y}{\nu_2 + \nu_1 y}\right)^{\frac{\nu_2}{2}-1}\frac{\nu_1 \nu_2}{(\nu_2 + \nu_1 y)^2}$$

$$= \frac{1}{B\left(\frac{\nu_1}{2}, \frac{\nu_2}{2}\right)} \frac{\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} y^{\frac{\nu_1}{2}-1}}{\left(1 + \frac{\nu_1}{\nu_2} y\right)^{\frac{\nu_1+\nu_2}{2}}} = \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} y^{\frac{\nu_1}{2}-1}}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)\left(1 + \frac{\nu_1}{\nu_2} y\right)^{\frac{\nu_1+\nu_2}{2}}}$$

$0 < x < 1 \Rightarrow y = \dfrac{\nu_2 x}{\nu_1(1-x)}$ olmak üzere $y > 0$ olur.

Bu durumda, Y r.d.'nin dağılımı,

$$f(y) = \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} y^{\frac{\nu_1}{2}-1}}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)\left(1 + \frac{\nu_1}{\nu_2} y\right)^{\frac{\nu_1+\nu_2}{2}}}, \quad y > 0$$

olarak elde edilir. Bu da F_{ν_1,ν_2} dağılımının olasılık yoğunluk fonksiyonudur.

ν_1 ve ν_2 serbestlik dereceleri ile F dağılımına sahip bir X raslantı değişkeni $(X \sim F_{\nu_1,\nu_2})$ için,

- X raslantı değişkeninin beklenen değeri $E(X) = \dfrac{\nu_2}{\nu_2 - 2}$ $(\nu_2 > 2 \text{ için})$'dir.
- X raslantı değişkeninin varyansı, $V(X) = \dfrac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2-2)^2(\nu_2-4)}$ $(\nu_2 > 4 \text{ için})$'dir.
- X raslantı değişkeninin orijine göre k'ıncı momenti,

$$\mu_k' = \left(\frac{\nu_2}{\nu_1}\right)^k \frac{\Gamma\left(\frac{\nu_1}{2}+k\right)}{\Gamma\left(\frac{\nu_1}{2}\right)} \frac{\Gamma\left(\frac{\nu_2}{2}-k\right)}{\Gamma\left(\frac{\nu_2}{2}\right)}, \quad \nu_2 > 2k$$

- F dağılımının moment çıkaran fonksiyonu yoktur. Bunun nedeni F dağılımının orijine göre k'ıncı momentinin yalnızca $k < \frac{\nu_2}{2}$ durumunda tanımlı olmasıdır.

TEOREM: X_1 ve X_2, sırasıyla v_1 ve v_2 parametreleriyle ki-kare dağılımına sahip bağımsız raslantı değişkenleri ise,

$$Y = \frac{X_1/v_1}{X_2/v_2}$$

biçiminde tanımlanan raslantı değişkeni F_{v_1,v_2} dağılımına sahip olur.

TEOREM: $X_1, X_2, \dots, X_{n_1}$, $N(\mu_1, \sigma_1^2)$ dağılımından rasgele örneklem; $Y_1, Y_2, \dots, Y_{n_2}$, $N(\mu_2, \sigma_2^2)$ dağılımından rasgele örneklem olsun. Bu örneklemlerin varyansları sırasıyla, S_1^2 ve S_2^2 olmak üzere,

$$\frac{S_1^2}{S_2^2} \sim F_{n_1-1,n_2-1}$$

yazılabilir.

İspat: Ki-kare Teorem 5'ten, $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ olduğu biliniyor.

Buna göre, $X_1, X_2, \dots, X_{n_1}$ örneklemi için, $\frac{(n_1-1)S_1^2}{\sigma_1^2} \sim \chi^2_{n_1-1}$

ve $Y_1, Y_2, \dots, Y_{n_2}$ örneklemi için, $\frac{(n_2-1)S_2^2}{\sigma_2^2} \sim \chi^2_{n_2-1}$

yazılabilir. Sonuç olarak,

$$\frac{\dfrac{\frac{(n_1-1)S_1^2}{\sigma_1^2}}{n_1-1}}{\dfrac{\frac{(n_2-1)S_2^2}{\sigma_2^2}}{n_2-1}} = \frac{S_1^2}{S_2^2} \sim F_{n_1-1,n_2-1}$$

olur.