İST 304 İstatistik Karar Kuramı ve Yöntemleri **Hafta XIII**

Ders 1 : Sonsal Kayıplarla Bayes Tahmin Edicisinin Elde Edilmesi

Dersi anlatan: İhsan Karabulut *Notları yazan: İhsan Karabulut*

Not: *LaTeX ders kalıbı UC Berkeley EECS Bölümü'nündür.*

Parametrelerin minimaks ve Bayes tahmin edicileri de veri olduğunda karar fonksiyonlarının seçiminde izlenen yol kullanılarak elde edilir. Parametre tahminine konu olan Θ doğa durumları uzayı (parametre uzayı) pek çok tahmin probleminde ikiden fazla eleman içerdiğinden grafik yöntemi faydalanılamaz bir araç olmaktadır. Hatırlanacağı gibi Bayes karar fonksiyonlarının elde edilmesi sürecinde $L_h(d(x_j), x_j)$ beklenen sonsal kayıpların kullanımıyla bu zorluk aşılmıştı. Torbada 5 bilye örneği, Bayes tahmin edicilerinin beklenen sonsal kayıpların kullanımıyla elde edilmesine örnek olarak verilecektir. $T = t(X)$ verilen bir parametre tahmin edicisi olmak üzere Bayes kaybını daha önce herhangi bir d karar fonksiyonu için yazılmış olan ve bunu beklenen sonsal kayıplarla olan ifadeye taşıyan eşitlik parametre tahmin edicisi için

$$\begin{aligned}
B(T) &= \sum_{i=1}^{m} R(\theta_i, d) g(\theta_i) \\
&= \sum_{j=1}^{k} \sum_{i=1}^{m} l(\theta_i, t(x_j)) h(\theta_i \,|x_j\,) P(x_j) \\
&= \sum_{j=1}^{k} \sum_{i=1}^{m} \{ l(\theta_i, t(x_j)) h(\theta_i \,|x_j\,) \} P(x_j) \\
&= \sum_{j=1}^{k} L_h(t(x_j), x_j) P(x_j)
\end{aligned}$$

dır. Her x_j gözlemi için en küçük $L_h(t(x_j), x_j)$ belenen sonsal kaybın seçimi ile Bayes kaybı en küçük olan T_B Bayes tahmin edicisine ulaşılacaktır, tüm T tahmin edicileri için $B(T_B) \leq B(T)$ olacaktır.

Örnek(Torbada 5 bilye örneğine devam). Beyaz bilye sayısının Bayes tahmin edicisi T_B elde edilecektir. Herhangi bir tahmin edici T ile gösterilecek olup parametre uzayı - doğa durumları uzayı - için verilen önsel dağılım

$g(\theta)$

$\boldsymbol{\theta} = \theta_i$	**0**	**1**	**2**	**3**	**4**	**5**
$P(\boldsymbol{\theta} = \theta)$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{2}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

olarak seçilmiştir. Kayıp fonksiyonu da daha önce verildiği gibi $l(\theta_i, a_j) = (a_j - \theta_i)^2$ karesel olsun.

Sırasıyla X'in marjinal olasılık fonksiyonu $P(x) = P(X = x)$, her x_j gözlemi koşulu altında $h(\theta_i \,|x_j\,)$ sonsal dağılımlar ve herhangi bir tahmin edicinin bir x_j gözlemi için beklenen kayıp $L_h(t(x_j), x_j) = L_h(a_i, x_j)$, $i = 1, 2, \cdots, 6$ değerleri elde edilecektir.

X r.d.'nin marjinal dağılımı için $P(X = 0)$ olasılığı

$$
\begin{aligned}
P(X=0) &= \sum_{i=1}^{5} P(X=0, \boldsymbol{\theta} = \theta_i) \\
&= \sum_{i=1}^{5} P(X=0 \,|\boldsymbol{\theta} = \theta_i\,) P(\boldsymbol{\theta} = \theta_i) \\
&= \sum_{i=1}^{5} \frac{(5-\theta_i)(4-\theta_i)}{20} P(\boldsymbol{\theta} = \theta_i) \\
&= \frac{(5-0)(4-0)}{20} \times \frac{1}{10} + \frac{(5-1)(4-1)}{20} \times \frac{2}{10} + \frac{(5-2)(4-2)}{20} \times \frac{3}{10} \\
&+ \frac{(5-3)(4-3)}{20} \times \frac{2}{10} + \frac{(5-4)(4-4)}{20} \times \frac{1}{10} + \frac{(5-5)(4-5)}{20} \times \frac{1}{10} \\
&= \frac{33}{100}
\end{aligned}
$$

olarak elde edilir. Benzer olarak $P(X = 1)$ ve $P(X = 2)$ olasılıkları da hesaplandığında marjinal dağılım olasılık dağılımı

$P\,(x\,)$

$X=x$	**0**	**1**	**2**
$P(X{=}x)$	$\frac{33}{100}$	$\frac{42}{100}$	$\frac{25}{100}$

olarak elde edilir. Aşağıda tüm $X = x_j$ gözlem değerleri koşulu altında elde edilmiş $h(\theta|x_j)$ sonsal dağılımları verilmiştir. Aşağıda sadece $h(\theta_2|1) = h(1|1)$ olasılığının hesaplanması verilecektir:

$$
\begin{aligned}
P(\boldsymbol{\theta} = 1\,|X = 1\,) &= h(1|1) \\
&= \frac{P(\boldsymbol{\theta} = 1, X = 1)}{P(X = 1)} \\
&= \frac{P(X = 1\,|\boldsymbol{\theta} = 1\,)P(\boldsymbol{\theta} = 1)}{P(X = 1)} \\
&= \frac{\frac{2\times(5-2)}{20} \times \frac{2}{10}}{\frac{42}{100}} \\
&= \frac{4}{21}
\end{aligned}
$$

$h\,(\theta\,|0\,)$

$\boldsymbol{\theta}=\theta_i$	**0**	**1**	**2**	**3**	**4**	**5**
$P(\boldsymbol{\theta}{=}\theta\vert X{=}0)$	$\frac{10}{33}$	$\frac{12}{33}$	$\frac{9}{33}$	$\frac{2}{33}$	0	0

$h\,(\theta\,|1\,)$

$\boldsymbol{\theta}=\theta_i$	**0**	**1**	**2**	**3**	**4**	**5**
$P(\boldsymbol{\theta}{=}\theta\vert X{=}1)$	0	$\frac{4}{21}$	$\frac{9}{21}$	$\frac{6}{21}$	$\frac{2}{21}$	0

$h\,(\theta\,|2\,)$

$\boldsymbol{\theta}=\theta_i$	**0**	**1**	**2**	**3**	**4**	**5**
$P(\boldsymbol{\theta}{=}\theta\vert X{=}2)$	0	0	$\frac{3}{25}$	$\frac{6}{25}$	$\frac{6}{25}$	$\frac{10}{25}$

Verilen sonsal kayıplardan yalnızca $L_h = (t(x_j), x_j) = L_h(a_3, 1)$'nın hesaplanması verilecektir. Hesaplar yapılırken $l(\theta_i, a_j) = (a_j - \theta_i)^2$ kayıp fonksiyonu için daha önce verilen kayıp fonksiyonu tablosu kullanılacaktır. $L_h(a_3, 1)$ sonsal kaybı gösterimininde a_3 eylemi için $\hat{\theta} = a_3 = \theta_3$ olduğu, bu nedenle $L_h(\theta_3, 1)$ olarak gösteriminin de kullanılabileceği hatırlanmalıdır. $X = 1$ gözlemi

$$L_h(t(x_j),x_j) = L_h(a_l,x_j)$$

	a_1	a_2	a_3	a_4	a_5	a_6
$X=0$	$\frac{22}{11}$	$\frac{9}{11}$	$\frac{18}{11}$	$\frac{49}{11}$	$\frac{102}{11}$	$\frac{177}{11}$
$X=1$	$\frac{42}{7}$	$\frac{17}{7}$	$\frac{6}{7}$	$\frac{9}{7}$	$\frac{26}{7}$	$\frac{57}{7}$
$X=2$	$\frac{412}{25}$	$\frac{241}{25}$	$\frac{120}{25}$	$\frac{49}{25}$	$\frac{28}{25}$	$\frac{57}{25}$

yapıldığında bilinmeyen θ parametresi için $t(1)=\theta_3=2$ tahminin yapılması halinde sonsal kayıp

$$\begin{aligned}
L_h(t(1),1) &= L_h(a_3,1)\\
&= L_h(\theta_3,1)\\
&= l(a_3,\theta_1)\times h(\theta_1\,|1) + l(a_3,\theta_1)\times h(\theta_1\,|1) + l(a_3,\theta_1)\times h(\theta_1\,|1)\\
&+ l(a_3,\theta_1)\times h(\theta_1\,|1) + l(a_3,\theta_1)\times h(\theta_1\,|1) + l(a_3,\theta_1)\times h(\theta_1\,|1)\\
&= 4\times 0 + 1\times\frac{4}{21} + 0\times\frac{9}{21} + 1\times\frac{6}{21} + 4\times\frac{2}{21} + 9\times 0\\
&= \frac{6}{7}
\end{aligned}$$

olarak hesaplanacaktır.

Yukarıda herhangi bir parametre tahmin edicisi T'nin Bayes kaybı $B(T)=\sum_{j=1}^{k} L_h(t(x_j),x_j)P(x_j)$ olarak ifade edilmişti. $B(T)$'nin en küçük olmasının her bir x_j gözlem değeri için $L_h(t(x_j),x_j)$ değerlerinin en küçüklerinin belirlenmesiyle olanaklı olduğu sonucuna varılmıştı. Yukarıdaki sonsal kayıplar tablosundan da anlaşılabileceği gibi $l=1,2,3,4,5,6$ için

$$\begin{aligned}
L_h(a_2,0) &\leq L_h(a_l,0)\\
L_h(a_3,1) &\leq L_h(a_l,1)\\
L_h(a_5,2) &\leq L_h(a_l,2)
\end{aligned}$$

dır.

Böylece $X = 0$ gözleminin yapılması halinde $a_2 = \theta_2 = 1$, $X = 1$ gözleminin yapılması halinde $a_3 = \theta_3 = 2$ ve $X = 2$ gözleminin yapılması halinde de $a_3 = \theta_5 = 4$ tahmininin yapılmasının sonsal kayıpları en az yapacağı görülür. O halde bilinmeyen θ parametresinin verilen önsel dağılım altında Bayes tahmin edicisi

$$T_B = \begin{cases} 1 & , X = 0 \\ 2 & , X = 1 \\ 4 & , X = 2 \end{cases}$$

dir. Bu tahmin edicinin Bayes riski

$$\begin{aligned} B(T_B) &= \sum_{j=1}^{k} L_h(t_B(x_j), x_j) P(x_j) \\ &= L_h(t_B(0), 0) \times P(X = 0) + L_h(t_B(1), 1) \times P(X = 1) \\ &\quad + L_h(t_B(2), 2) \times P(X = 0) \\ &= L_h(a_2, 0) \times P(X = 0) + L_h(t_B(a_3), 1) \times P(X = 1) \\ &\quad + L_h(t_B(a_5), 2) \times P(X = 0) \\ &= \frac{9}{11} \times \frac{33}{100} + \frac{6}{7} \times \frac{42}{100} + \frac{28}{25} \times \frac{25}{100} \\ &= \frac{91}{100} \end{aligned}$$

dir. T_B tahmin edicisinin risk fonksiyonu her $\theta \in \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6\}$ için

$$
\begin{aligned}
R(\theta, t_B) &= E(l(\theta, t_B(X))^2) \\
&= \sum_{x=0}^{3} l(\theta, t_B(x_j))^2 P_\theta(X = x) \\
&= (t_B(0) - \theta)^2 \frac{(5-\theta)(4-\theta)}{20} + (t_B(1) - \theta)^2 \frac{2\theta(5-\theta)}{20} \\
&\quad + (t_B(2) - \theta)^2 \frac{\theta(\theta-1)}{20} \\
&= \frac{(1-\theta)^2(5-\theta)(4-\theta)}{20} + \frac{(2-\theta)^2 2\theta(5-\theta)}{20} \\
&\quad + \frac{(4-\theta)^2\theta(\theta-1)}{20} \\
&= 1 - \frac{5\theta}{4} + \frac{3\theta^2}{4} - \frac{\theta^3}{10}
\end{aligned}
$$

dır. Karesel kayıp fonksiyonu kullanıldığı için tahmin edicinin $HKO(T_B)$' ye eşittir. Tahmin edicinin beklenen değeri

$$
\begin{aligned}
E(T_B) &= \sum_{x=0}^{3} t_B(x) P_\theta(X = x) \\
&= 1 \times \frac{(5-\theta)(4-\theta)}{20} + 2 \times \frac{2\theta(5-\theta)}{20} + 4 \times \frac{\theta(\theta-1)}{20} \\
&= 1 + \frac{7\theta}{20} + \frac{\theta^2}{20}
\end{aligned}
$$

olup yanlı bir tahmin edicidir. Yanın karesi

$$
b^2(T_B) = \frac{1}{400}(20 - 13\theta + \theta^2)
$$

Değerlendirme yapılırken örneklem çapının yalnızca $n = 2$ olduğu hatırlanmalıdır.

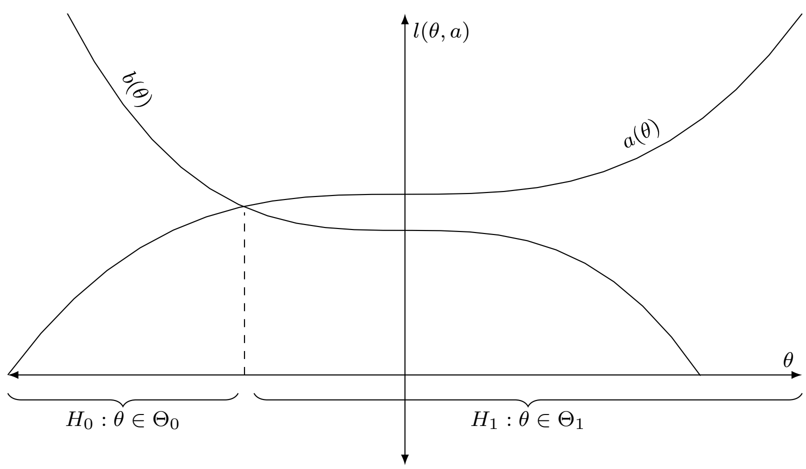
İST 304 İstatistik Karar Kuramı ve Yöntemleri **Hafta XIII**

Ders 2 : Karar Verme Problemi Olarak İstatistiksel Hipotezlerin Testi

Dersi anlatan: İhsan Karabulut *Notları yazan: İhsan Karabulut*

İstatistiksel hipotezlerin testi sadece iki eylemin yer aldığı tipik bir karar verme problemi olarak ele alınabilir. İstatiksel hipotezlerin testi kapsamlı bir içeriğe sahip olmasına karşın bu alt başlıkta sadece dağılım parametrelerine ilişkin basit hipotezlerin testi konusu ele alınacaktır. İstatistiksel hipotezlerin testi de önceki konularda olduğu gibi öncelikle rasgele gözlem kullanılmadan ele alınabilirdi ancak buna gerek duyulmadı ve rasgele gözlemlere dayalı olarak verilmeye çalışıldı. İstatistiksel hipotezlerin testinde de gözlemi yapılan rasgele değişkenlerin doğa durumları ile ilgili olduğu düşünülecektir. Bu nedenle doğa durumlarının uzayı parametre uzayı Θ olacaktır. Eylem uzayı $\mathcal{A} = \{A, B\}$ olarak gösterilecektir. Eylemler rasgele değişken gözlemleri üzerinde tanımlanan karar fonksiyonları ile belirleneceğinden eylemler rasgele olaylar (H_0'ı reddedememek veya H_1'i kabul etmek) olacaktır. A eylemi daha önce değinildiği gibi H_0 yokluk hipotezini reddedememek anlamında H_0'ı kabul etmek eylemini gösterecektir, bu eylemin gerçekleşmesi halinde rasgele değişken 0 değerini alacaktır; B, H_1 alternatif hipotezini kabul etmek eylemini gösterecektir, B eyleminin gerçekleşmesi halinde rasgele değişken 1 değerini alacaktır. Θ da yer alan elemanlardan yalnızca bir θ'nın doğanın durumu olduğunun ifade edilmesi **basit hipotez** olarak adlandırılır. Hipotez testine konu olan doğanın durumu gerçekte θ ise bu hipotez doğrudur denilir. Doğanın durumunun birden fazla θ dan biri olabileceğinin ifade edildiği hipotez ise **karmaşık (composite) hipotez** olarak adlandırılır. Doğanın gerçek durumu karmaşık hipotezde içerilen doğa durumlarından biri ise- hipotez testine konu olan parametresi bunlardan biri ise- hipotez doğrudur denilir. $\Theta = \Theta_0 \bigcup \Theta_1$, $\Theta_0 \bigcap \Theta_1 = \emptyset$ olmak üzere dağılım parametresine ilişkin yokluk ve alternatif hipotezler sırasıyla $H_0 : \theta \in \Theta_0$, $H_1 : \theta \in \Theta_1$ olarak ifade edilirler. H_0 ve H_1 basit hipotezler ise söz konusu karar problemi doğanın $\Theta = \{\theta_0, \theta_1\}$, $\Theta_0 = \{\theta_0\}$, $\Theta_1 = \{\theta_1\}$ olan iki durumunun olduğu ve iki eylemin olduğu bir karar problemidir. Rasgele gözlemlerin olması halinde karar fonksiyonu test istatistiği olacaktır. Bir karar probleminin tanımlı olabilmesi için kayıp fonksiyonu da tanımlı olmalıdır.

Kayıp fonksiyonu Θ herhangi bir parametre uzayı ve $a \in \mathcal{A} = \{A, B\}$ eylemi için $l(\theta, a)$ kaybı θ'nın bir fonksiyonu olacaktır ve karar vericinin yapabileceği iki eylem ile tanımlanacaktır: A eylemi için $l(\theta, A) = a(\theta)$ ve B için $l(\theta, B) = b(\theta)$. Kayıp fonksiyonu verilecek yanlış kararı kayıpla sonuçlandıracak şekilde tanımlanmalıdır. Dağılım parametresi için tipik bir hipotez testinde bu durum Şekil' de yansıtılmıştır. H_0 ın doğruluğu bilindiğinde doğru eylem A olacak, B eylemine göre kayıp daha az; H_1in doğruluğu bilindiğinde doğru eylem B olacak, A eylemine göre kayıp daha az olmalıdır. Her eylem için kayıp fonksiyonunu doğa durum uzayını iki kümeye ayıracaktır: θ'nın H_0 hipotezi ile ortaya konulan kümede yer alması halinde her $\theta \in \Theta_0$ parametre değeri için H_0'ı kabul etme eylemi için $a(\theta) \leq b(\theta)$ olmalıdır, benzer olarak her $\theta \in \Theta_1$ parametre değeri için H_1'ı kabul etme eylemi için $b(\theta) \leq a(\theta)$ olmalıdır.



Şekil 2.1: İstatistiksel hipotez testlerinde kullanılabilecek kayıp fonksiyonları ve işlevleri.

Bayes ilkesinin kullanıldığı karar verme problemlerinde kayıp ya da pişmanlık fonksiyonunu kullanılmasıyla verilecek kararın değişmediği biliniyor. Bu nedenle Bayes ilkesinin kullanılacağı hipotezlerin testinde kayıp fonksiyonu yerine pişmanlık fonksiyonu ile çalışmak değerlendirme yapmayı kolaylaştıracaktır. **Pişmanlık fonksiyonunun oluşturulmasında** aşağıdaki kayıp fonksiyonu tablosu dikkate alınsın:

$l(\theta, a)$

	A	B
$H_0 : \theta \in \Theta_0$	$l(\theta, A)$	$l(\theta, B)$
$H_1 : \theta \in \Theta_1$	$l(\theta, A)$	$l(\theta, B)$

Pişmanlık fonksiyonunun tanımı ve $\theta \in \Theta_0$ için $l(\theta, A) \leq l(\theta, B)$ ve $\theta \in \Theta_1$ için $l(\theta, B) \leq l(\theta, A)$ olduğu kullanılırsa pişmanlık fonksiyonu aşağıdaki gibi elde edilecektir:

$r(\theta, a)$

	A	B
$H_0 : \theta \in \Theta_0$	0	$l(\theta, B) - l(\theta, A)$
$H_1 : \theta \in \Theta_1$	$l(\theta, A) - l(\theta, B)$	0

Yalnızca iki sade eylem olduğundan

$$r(\theta, A) = \begin{cases} 0 & , H_0 \text{ doğru} \\ a(\theta) - b(\theta) & , H_1 \text{ doğru} \end{cases}$$

ve

$$r(\theta, B) = \begin{cases} b(\theta) - a(\theta) & , H_0 \text{ doğru} \\ 0 & , H_1 \text{ doğru} \end{cases}$$

olarak da yazılabilir.

Pişmanlık fonksiyonu daha da basit yazılabilir. $\theta \in \Theta_0$ için $l(\theta, A) = 0$, $\theta \in \Theta_1$ için $l(\theta, B) = 0$, ve $a \in \mathcal{A} = \{A, B\}$ olmak üzere her θ için $l(\theta, a) \geq 0$ olarak tanımlanmış ve $\theta \in \Theta_0$ için $l(\theta, B) = b(\theta)$

ve $\theta \in \Theta_1$ için $l(\theta, A) = a(\theta)$ ise $r(\theta, a)$ aşağıdaki tabloda verildiği gibi olacaktır:

$r(\theta, a)$	A	B
$H_0 : \theta \in \Theta_0$	0	$b(\theta)$
$H_1 : \theta \in \Theta_1$	$a(\theta)$	0

Hipotezler hemen sonra ele alacağımız gibi hipotezler $H_0 : \theta = \theta_0$, $H_1 : \theta = \theta_1$ basitse $a > 0$ ve $b > 0$ olmak üzere pişmanlık fonksiyonu aşağıdaki gibi yazılabilir.

$r(\theta, a)$	A	B
$H_0 : \theta = \theta_0$	0	b
$H_1 : \theta = \theta_1$	a	0

Buraya kadar istatistiksel hipotez testlerinden bildik olan $I.tip$ α ve $II.tip$ β hatalarından söz edilmedi. Bu hataların pişmanlık fonksiyonu ile tanımlanacağı ileride görülecektir. $I.tip$ α hatasının gerçekte doğru olan H_0 hipotezinin kabul edilmemesiyle (B eylemiyle) yaşanacak pişmanlık değeri b ile ve $II.tip$ β hatasının gerçekte doğru olan H_1 hipotezinin kabul edilmemesiyle (A eylemiyle) yaşanacak pişmanlık değeri a ile ilgili olduğu anlaşılacaktır.

Açıklama. Bu ifadelede daha önceki bildik "... verilen α anlamlılık düzeyinde... " ifadesi yoktur. Hipotezlerin testi konusunda istatistiksel karar teorisinde, klasik istatistik teorisinde olmayan bir simetri olduğu görülür; karar teorisinde her iki hipotez de aynı rollere sahiptir. Klasik istatistik teorisinde her iki hipotez aynı rollere sahip değildirler, test H_0 hipotezinin doğru olduğu varsayımıyla gerçekleştirilir.

Parametre değerlerine ilişkin basit hipotezlerin istatistiksel testinin bir karar problemi olarak tanımlanmasını sağlayacak unsurlar tanımlanmış bulunuyor. Basit hipo- tezlerin istatistiksel testi bir karar problemi olarak ele alınabilir. Gözlemlerin hipotezlerin test edilmesindeki karar sürecine katılması karar fonksiyonlarının uygun olarak tanımlanmasıyla olanaklı dır.

Tanım. X rasgele gözlemlerine dayalı olarak H_0ın red edilmesi eylemi B'ye yol açan gözlemler kümesi **testin kritik bölgesi** olarak adlandırılır. Testin kritik bölgesi C ile gösterilecektir.

Tanım. X rasgele gözlem değerleri üzerinde tanımlı karar fonksiyonu **test istatistiği** olarak adlandırılır. Test istatistiği $T(X)$ ile gösterilecektir.

Kritik bölge ve test istatistiğinden biri diğerini karakterize eder. C kümesi biliniyorsa test istatistiği; test istatistiği biliniyorsa C kümesi biliniyor demektir.

X rasgele değişkeninin alabileceği sonlu m sayıda değer varsa yalnızca iki sade eylemden biri A ya da B yapılabileceğinden 2^m tane (sade) karar fonksiyonu tanım-lanabilir. Bu karar fonksiyonlarından her biri H_0'ın H_1'e karşı testinde kullanılabilir.

Örnek. Doğa durumları uzayı $\Theta = \{\theta_0, \theta_1\}$ olmak üzere, doğa durumları ile ilgili olduğu düşünülen X rasgele değişkeni $x_1, x_2, \cdots, x_m$ değerlerini alsın ve $f(x, \theta) = P_\theta(X = x)$ olasılık fonksiyonuna sahip olsun:

$$f(x, \theta) = P_\theta(X = x)$$

$X = x_i$	x_1	x_2	$\cdots$	x_m
$P_{\theta_0}(X = x_i)$	p_1	p_2	$\cdots$	p_m
$P_{\theta_1}(X = x_i)$	r_1	r_2	$\cdots$	r_m

$H_0 : \theta = \theta_0$ hipotezi $H_1 : \theta = \theta_1$ hipotezine karşı test edilecektir. Her doğa durumunda sade eylemlerin pişmanlığını gösteren pişmanlık fonksiyonu tablosu aşağıdaki gibidir:

$r(\theta, a)$

	A	B
$H_0 : \theta = \theta_0$	0	1
$H_1 : \theta = \theta_1$	2	0

Herhangi bir test istatistiği (karar fonksiyonu) gözlem değerlerinin herhangi bir alt kümesi $\{x_i\} \subseteq \{x_1, x_2, \cdots, x_m\}$ olmak üzere

$$T(X) = \begin{cases} 1 & , X \in \{x_i\} \\ 0 & , X \notin \{x_i\} \end{cases}$$

dir. Burada kritik küme $C = \{x_i\}$ dir. X rasgele değişkeninin (istatistiğinin) gözlem değeri $\{x_i\}$ kümesi içinde yer aldığında B eylemi, gözlem değeri $\{x_i\}$ kümesi içinde yer almadığında A eylemi yapılacaktır. Yapılacak eylem rasgele gözleme göre yapıldığından yapılacak eylem de rasgele olacaktır; eylem B ise rasgele değişken olarak $T(X)$ istatistiği 1 değerini aksi halde 0 değerini alacaktır. Burada $C = \{x_i\}$ alt kümesinin veya $\{x_1, x_2, \cdots, x_m\}$ olabileceği hatırlanmalıdır.

Not. Gözlemi yapılarak hipotez testi gerçekleştirilecek olan rasgele değişken daha önceki konularda olduğu gibi sonlu k sayıda değer alacaktır. Söz konusu X rasgele değişkeninin $\sum_{i=1}^{n} X_i^r$, $\bar{X}$, vb. bir **istatistiği** temsil ettiği, x_i gözlenen değerlerinin de bu istatistiğin değerleri olduğu görülecektir.

Verilen bir test ile ilgili olduğu kritik bölge C aynı kararlara işaret edecekler ve aynı işlevlere sahip olacaklardır. Bu nedenle bir test istatistiği, $T = t(X)$ ile buna karşılık gelen C kritik bölgesi için verilen bir $\theta \in \Theta$ doğa durumunda risk fonksiyonu (ya da beklenen pişmanlık) değerlerine hesaplanabilecek ve

$$R(\theta, t) = R(\theta, C)$$

olacaktır. Bu nedenle risk fonksiyonu için iki notasyondan biri anlaşılabilirlik gözetile- rek kullanılacaktır.

Bundan sonraki konularda Bayes ilkesinin kullanımına uygun olarak X rasgele değişkeninin θ parametreli olasılık fonksiyonu $P_\theta(X = x) = f(x;\theta)$ koşullu olasılık fonksiyonu olarak anlaşılacak ve $P_\theta(X = x) = P(X = x\,|\theta\,) = f(x\,|\theta\,)$ olarak kullanılacaktır.

Kaynaklar

[1] D. Blackwell and M.A. Girshick (1979), Theory of games and statistical decisions, *Dover Publications*, New York.

[2] H. Chernoff and L. E. Moses (1986), Elementary Decision Theory, *Dover Publications*, New York.

[3] B. W. Lindgren (1971), Elements of Decision Theory, *Macmillan Company*, New York.