## HAFTA 6

# TAHMİN EDİCİLERDE ARANAN ÖZELLİKLER

Bir kitleyi karekterize eden o kitlenin parametreleridir. Kitle hakkında herhangi bir istatistiki sonuç çıkarım aslında kitle parametreleri hakkında bir sonuç çıkarımdır. Bir rasgele değişkenin dağılım fonksiyonu (veya olasılık fonksiyonu) kitle parametrelerine bağlıdır. Kitlenin parametre sayısı bir tane olabildiği gibi, birden fazla da olabilir. Örneğin, normal dağılımın iki tane parametresi vardır. Kitlenin parametresi hakkında tahminde bulunabilmek için deneyler tekrarlanır. Tekrar sayısı arttıkça, parametre hakkında daha iyi bilgi elde edileceği beklenir. Ancak, deney sayısının fazla olması bazen imkansız olduğu gibi bazen de pahalı olabilir. Böyle durumlarda, küçük örneklemler ile parametre hakkında bilgi sahibi olmak istenebilir.

Bir parametre hakkında tahminde bulanmak için herhangi bir engel yoktur. Yani, bir parametre için değişik tahminler önerilebilir. En iyi denilebilecek bir tahminden daha iyi tahminler de yapılabilir. Dolayısı ile, "*iyi bir tahminden*" ne anlamamız gerektiği iyi tanımlanmalıdır. Bunun için, iyi bir tahminde aranan özelliklerin olması gerekir. Ancak bu özellikler kapsamında iyi bir tahminden söz edilebilir. Tahmin, önerilen bir tahmin edicinin değeri olduğundan, tahminden söz edilirken aslında tahmin ediciden söz edilmektedir. Bu bölümde, yaygın olarak kabul gören bazı özellikler incelenecektir.

### 7.1. Yeterlilik Özelliği (Sufficiency)

Bilinmeyen $\theta$ parametresi için yeterli bir tahmin edici, örneklem içinde parametre hakkındaki bilgiyi özetleyen tahmin edicidir. Yani, örneklem içinde parametre hakkında ne kadar bilgi varsa, hiçbir bilgi kaybı olmadan özetleyen bir tahmin edici, o parametre için yeterlidir.

$X_1, X_2, \ldots, X_n$ olasılık veya olasılık yoğunluk fonksiyonu $f(x;\theta)$ olan kitleden bir örneklem olsun. $T(X_1, X_2, \ldots, X_n) = T(\underset{\sim}{X})$ tahmin edicisi $\theta$ için yeterli ise, $\theta$ hakkında $X_1, X_2, \ldots, X_n$ örneklemine bağlı olarak elde edilen bir sonuç sadece $T(\underset{\sim}{X})$ tahmin edicisinin değerine bağlıdır. Yani, $\underset{\sim}{x} = (x_1, x_2, \ldots, x_n)'$ ve $\underset{\sim}{y} = (y_1, y_2, \ldots, y_n)'$ ayrı ayrı gözlemler ve $T(\underset{\sim}{x}) = T(\underset{\sim}{y})$ ise, $\theta$ hakkında $X$ ler veya $Y$ ler üzerinden yapılacak sonuç çıkarımlar aynıdır. $X_1, X_2, \ldots, X_n$ örneklemi için bazen $\underset{\sim}{X}$ gösterimi de kullanılacaktır.

**Tanım 7.1.1** $T(\underset{\sim}{X})$ verildiğinde, $\underset{\sim}{X}$ in koşullu dağılımı $\theta$ ya bağlı değilse, $T(\underset{\sim}{X})$'e $\theta$ için *yeterli bir tahmin edici* denir $\otimes$

$\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)'$ bir örneklem olmak üzere $P(\underset{\sim}{X} = \underset{\sim}{x} | T(\underset{\sim}{X}) = t)$ koşullu olasılığı (sürekli durum için $f_{\underset{\sim}{X} | T(\underset{\sim}{X}) = t}(\underset{\sim}{x} | t)$ koşullu olasılık yoğunluk fonksiyonu) $\theta$ ya bağlı değilse $T(\underset{\sim}{X})$, $\theta$ için yeterlidir. $T(\underset{\sim}{X})$ tahmin edicisi yerine bazen $T$ kullanılacaktır. Tanımdan da anlaşılacağı gibi, herhangi bir $\theta$ parametresi için yeterli istatistik tek değildir. $T$, $\theta$ için yeterli ise gözlenen $\underset{\sim}{X}$ ve $\underset{\sim}{Y}$ ler için $P_\theta(\underset{\sim}{X} = \underset{\sim}{x}) = P_\theta(\underset{\sim}{Y} = \underset{\sim}{x})$ dir. $T$ yeterli olduğundan $P(\underset{\sim}{X} = \underset{\sim}{x} | T = t) = P(\underset{\sim}{Y} = \underset{\sim}{x} | T = t)$ dir. Ayrıca $\{\underset{\sim}{X} = \underset{\sim}{x}\} \subset \{T(\underset{\sim}{X}) = T(\underset{\sim}{x})\}$ olup $P_\theta(\underset{\sim}{X} = \underset{\sim}{x}) = P_\theta(\underset{\sim}{Y} = \underset{\sim}{x})$ eşitliği

$$\begin{aligned}P_\theta(\underset{\sim}{X} = \underset{\sim}{x}) &= P_\theta(\underset{\sim}{X} = \underset{\sim}{x}, T = t) = P(\underset{\sim}{X} = \underset{\sim}{x} | T = t) P_\theta(T = t) \\ &= P_\theta(\underset{\sim}{Y} = \underset{\sim}{x} | T = t) P_\theta(T = t) = P_\theta(\underset{\sim}{Y} = \underset{\sim}{x}, T = t) = P_\theta(\underset{\sim}{Y} = \underset{\sim}{x})\end{aligned}$$

şeklinde elde edilir. Diğer taraftan, $P(\underset{\sim}{X} = \underset{\sim}{x} | T = t)$ koşullu olasılığı

$$P(\underset{\sim}{X} = \underset{\sim}{x} | T = t) = \frac{P_\theta(\underset{\sim}{X} = \underset{\sim}{x}\ ,\ T = t)}{P_\theta(T = t)} = \frac{P_\theta(\underset{\sim}{X} = \underset{\sim}{x})}{P_\theta(T = t)} = \frac{p(\underset{\sim}{x}; \theta)}{q(T(\underset{\sim}{x}); \theta)}$$

şeklinde yazılabilir. Buna göre, aşağıdaki teoremi ispatsız olarak verebiliriz.

**Teorem 7.1.1** $T$ tahmin edicisinin $\theta$ için yeterli olması için gerek ve yeter koşul $p(\underset{\sim}{x}; \theta) / q(T(\underset{\sim}{x}); \theta)$ oranının $\theta$ dan bağımsız olmasıdır ◊

**Örnek 7.1.1** a) $X_1, X_2, \ldots, X_n$ parametresi $p$ olan Bernoulli dağılımından bir örneklem olsun. $T = X_1 + X_2 + \ldots + X_n$ tahmin edicisinin $p$ için yeterli olup olmadığını araştıralım. $T \sim Binom(n, p)$ olup olasılık fonksiyonu,

$$P_p(T = t) = \binom{n}{t} p^t (1 - p)^{n-t} \ ,\ t = 0, 1, 2, \ldots, n$$

şeklindedir. $t = x_1 + x_2 + \ldots + x_n$ için $p(\underset{\sim}{x}; p) / q(T(\underset{\sim}{x}); p)$ oranı $p$ ye bağlı değilse $T$, $p$ için yeterlidir. Teorem (7.1.1) e göre $t = x_1 + x_2 + \ldots + x_n$ için bu oran,

$$\frac{p(\underset{\sim}{x}; p)}{q(T(\underset{\sim}{x}); p)} = \frac{P_p(\underset{\sim}{X} = \underset{\sim}{x})}{P_p(T = t)} = \frac{\prod_{i=1}^{n} P_p(X_i = x_i)}{P_p(T = t)} = \frac{\prod_{i=1}^{n} p^{x_i} (1 - p)^{1 - x_i}}{\binom{n}{t} p^t (1 - p)^{n-t}} = \frac{p^{\sum_{i=1}^{n} x_i} (1 - p)^{n - \sum_{i=1}^{n} x_i}}{\binom{n}{t} p^t (1 - p)^{n-t}} = \frac{1}{\binom{n}{t}}$$

şeklinde olup $p$ ye bağlı değildir. $T = \sum_{i=1}^{n} X_i$ tahmin edicisi $p$ parametresi için yeterlidir.

b) $X_1, X_2, \ldots, X_n$ beklenen değeri $\mu$ olan normal dağılımdan bir örneklem olsun. $T = X_1 + X_2 + \ldots + X_n \sim N(n\mu, n)$ olup, $T$ nin $\mu$ için yeterli olup olmadığını inceleyelim. Bunun için, Teorem (7.1.1) deki oranın payı

$$p(\underset{\sim}{x};\mu) = \prod_{i=1}^{n} f_\mu(x_i;\mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x_i-\mu)^2\right) = \left(\frac{1}{2\pi}\right)^{n/2} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

$$= \left(\frac{1}{2\pi}\right)^{n/2} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2 + \mu\sum_{i=1}^{n} x_i - \frac{n\mu^2}{2}\right)$$

ve paydası

$$q(T(\underset{\sim}{x});\mu) = q(t;\mu) = \frac{1}{\sqrt{2n\pi}} \exp\left(-\frac{1}{2n}(t-n\mu)^2\right) = \frac{1}{\sqrt{2n\pi}} \exp\left(-\frac{1}{2n}(t^2 - 2n\mu t + n^2\mu^2\right)$$

$$= \frac{1}{\sqrt{2n\pi}} \exp\left(-\frac{t^2}{2n} - \frac{n^2\mu^2}{2n} + \mu t\right)$$

dir. $t = x_1 + x_2 + \ldots + x_n$ için oran

$$\frac{p(\underset{\sim}{x};\mu)}{q(T(\underset{\sim}{x});\mu)} = \frac{p(\underset{\sim}{x};\mu)}{q(t;\mu)} = \frac{\left(\frac{1}{2\pi}\right)^{n/2} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2 + \mu\sum_{i=1}^{n} x_i - \frac{n\mu^2}{2}\right)}{\frac{1}{\sqrt{2n\pi}} \exp\left(-\frac{t^2}{2n} + \mu t - \frac{n\mu^2}{2}\right)}$$

$$= \frac{\sqrt{2n\pi}}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}\left[\sum_{i=1}^{n} x_i^2 - \frac{t^2}{n}\right]\right)$$

şeklinde yazılır. Bu oran $\mu$ ye bağlı olmadığından $T$, $\mu$ için yeterlidir.

c) Benzer şekilde, $X_1, X_2, \ldots, X_n$ parametresi $\theta$ olan Poisson dağılımından bir örneklem olsun. $T = X_1 + X_2 + \ldots + X_n \sim Poisson(n\theta)$ olup,

$$\frac{p(\underset{\sim}{x};\theta)}{q(T(\underset{\sim}{x});\theta)} = \frac{P_\theta(\underset{\sim}{X} = \underset{\sim}{x})}{P_\theta(T=t)} = \frac{\prod_{i=1}^{n} P_\theta(X_i = x_i)}{P_\theta(T=t)} = \frac{e^{-n\theta}\, \theta^{\sum_{i=1}^{n} x_i} \prod_{i=1}^{n} \frac{1}{x_i!}}{e^{-n\theta}\,(n\theta)^t / t!} = \frac{t!}{n^t} \prod_{i=1}^{n} \frac{1}{x_i!}$$

oranı $t = x_1 + x_2 + \ldots + x_n$ için $\theta$ parametresinden bağımsızdır. Yani $T$, $\theta$ için yeterlidir $\oplus$

Tanımdan hareketle yeterli istatistiğin bulunması bazen zor olabilir. Bir tahmin edici verildiğinde, onun yeterli olup olmadığı da araştırılabilir. Ancak yeterli tahmin ediciyi önceden kestirmek kolay olmayabilir. Aşağıdaki teorem yeterli istatistiği bulmak için önemli bir araçtır.

**Teorem 7.1.2** *(Faktörizasyon Teoremi)* $X_1, X_2, \ldots, X_n$ olasılık veya olasılık yoğunluk fonksiyonu $f(x;\theta)$ olan kitleden bir örneklem olsun. $T$, $\theta$ için yeterli olması için gerek ve yeter koşul öyle $g(t;\theta)$ ve $h(x)$ fonksiyonları için, ortak olasılık ve olasılık yoğunluk fonksiyonunun $f(\underset{\sim}{x};\theta) = g(T(\underset{\sim}{x});\theta)\,h(\underset{\sim}{x})$ şeklinde yazılabilmesidir. $g$ fonksiyonu $T(\underset{\sim}{x})$ ve $\theta$ nın bir fonksiyonu olup $h$ fonksiyonu ise sadece $\underset{\sim}{x}$ in bir fonksiyonudur.

*İspat:* Kesikli durumu gözönüne alalım (sürekli durumda ispat Hogg, and Craig (1995) sayfa 318 de verilmiştir). $T$ tahmin edicisi $\theta$ için yeterli olsun. Yeterliliğin tanımından $P(\underset{\sim}{X} = \underset{\sim}{x} | T = t)$ koşullu olasılığı $\theta$ ya bağlı değildir. Buradan, $h(\underset{\sim}{x}) = P(\underset{\sim}{X} = \underset{\sim}{x} | T = t)$ ve $g(T(\underset{\sim}{x});\theta) = P_\theta(T = t)$ fonksiyonları seçildiğinde, $f(\underset{\sim}{x};\theta)$ ortak olasılık fonksiyonu,

$$f(\underset{\sim}{x};\theta) = P_\theta(\underset{\sim}{X} = \underset{\sim}{x}) = P_\theta(T = t)P(\underset{\sim}{X} = \underset{\sim}{x} | T = t) = g(T(\underset{\sim}{x});\theta)\,h(\underset{\sim}{x})$$

şeklinde yazılabilir. Böylece, teoremin gerek kısmı ispatlanmış olur.

Şimdi, öyle $g(t;\theta)$ ve $h(x)$ fonksiyonları için ortak olasılık fonksiyonu

$$f(\underset{\sim}{x};\theta) = g(T(\underset{\sim}{x});\theta)\,h(\underset{\sim}{x})$$

şeklinde yazılabilsin. $p(\underset{\sim}{x};\theta)/q(T(\underset{\sim}{x});\theta)$ oranı $\theta$ ya bağlı değilse $T$ tahmin edicisi $\theta$ için yeterlidir. $A_{T(\underset{\sim}{x})} = \{\underset{\sim}{y} : T(\underset{\sim}{x}) = T(\underset{\sim}{y})\}$ olmak üzere $T(\underset{\sim}{x})$ fonksiyonu $A_{T(\underset{\sim}{x})}$ kümesi üzerinde sabittir. Ayrıca, $T$ nin olasılık fonksiyonu, $q(T(\underset{\sim}{x});\theta) = \sum_{\underset{\sim}{y} \in A_{T(\underset{\sim}{x})}} q(T(\underset{\sim}{y});\theta)\,h(\underset{\sim}{y})$ şeklinde yazılabildiğinden Teorem (7.1.1) deki oran,

$$\frac{p(\underset{\sim}{x};\theta)}{q(T(\underset{\sim}{x});\theta)} = \frac{g(T(\underset{\sim}{x});\theta)h(\underset{\sim}{x})}{q(T(\underset{\sim}{x});\theta)} = \frac{g(T(\underset{\sim}{x});\theta)h(\underset{\sim}{x})}{\sum_{\underset{\sim}{y} \in A_{T(\underset{\sim}{x})}} g\left(T(\underset{\sim}{y});\theta\right)h(\underset{\sim}{y})}$$

$$= \frac{g(T(\underset{\sim}{x});\theta)h(\underset{\sim}{x})}{g(T(\underset{\sim}{x});\theta) \sum_{\underset{\sim}{y} \in A_{T(\underset{\sim}{x})}} h(\underset{\sim}{y})} = \frac{h(\underset{\sim}{x})}{\sum_{\underset{\sim}{y} \in A_{T(\underset{\sim}{x})}} h(\underset{\sim}{y})}$$

olup parametreden bağımsızdır. Yani $T$, $\theta$ için yeterlidir ◊

**Örnek 7.1.2** a) $X_1, X_2, \ldots, X_n$ parametresi $p$ olan Bernoulli dağılımından bir örneklem olsun. $T = X_1 + X_2 + \ldots + X_n$ nin $p$ için yeterli olduğunu bir önceki örnekten biliyoruz. Şimdi, faktörizasyon teoremini kullanarak $T$ nin yeterli olduğunu görelim. $X$ lerin ortak olasılık fonksiyonu,

$$f(\underset{\sim}{x};p)=P_p\left(\underset{\sim}{X}=\underset{\sim}{x}\right)=\prod_{i=1}^{n}P_p(X_i=x_i)=\prod_{i=1}^{n}p^{x_i}(1-p)^{1-x_i}$$

$$=p^{\sum_{i=1}^{n}x_i}(1-p)^{n-\sum_{i=1}^{n}x_i}=p^{T(\underset{\sim}{x})}(1-p)^{n-T(\underset{\sim}{x})}$$

şeklinde olup, ortak olasılık fonksiyonu $g(T(\underset{\sim}{x});p)=p^{T(\underset{\sim}{x})}(1-p)^{n-T(\underset{\sim}{x})}$ ve $h(\underset{\sim}{x})=1$ için

$$f(\underset{\sim}{x};p)=g(T(\underset{\sim}{x});p)\,h(\underset{\sim}{x})$$

şeklinde yazılabildiğinden faktörizasyon teoremine göre $T$ tahmin edicisi $p$ için yeterlidir.

b) $N(\mu,\sigma^2)$ dağılımından bir örneklem $X_1,X_2,\ldots,X_n$ olsun. $X_1,X_2,\ldots,X_n$ rasgele değişkenlerinin ortak olasılık yoğunluk fonksiyonu, $\underset{\sim}{T}\left(\underset{\sim}{x}\right)=\left(\sum_{i=1}^{n}x_i,\ \sum_{i=1}^{n}x_i^2\right)$ için,

$$f\left(\underset{\sim}{x};\mu,\sigma^2\right)=\left(\frac{1}{2\pi\sigma^2}\right)^{n/2}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}x_i^2+\frac{\mu}{\sigma^2}\sum_{i=1}^{n}x_i-\frac{n\mu^2}{2\sigma^2}\right)$$

$$=\left(\frac{1}{2\pi\sigma^2}\right)^{n/2}\exp\left(-\frac{n\mu^2}{2\sigma^2}\right)\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}x_i+\frac{\mu}{\sigma^2}\sum_{i=1}^{n}x_i\right)$$

$$=\left(\frac{1}{2\pi}\right)^{n/2}\left(\frac{1}{\sigma^2}\right)^{n/2}\exp\left(-\frac{n\mu^2}{2\sigma^2}\right)\exp\left[\left(\frac{\mu}{\sigma^2},-\frac{1}{2\sigma^2}\right)\begin{pmatrix}\sum_{i=1}^{n}x_i\\ \sum_{i=1}^{n}x_i\end{pmatrix}\right]$$

$$=\left(\frac{1}{2\pi}\right)^{n/2}\left(\frac{1}{\sigma^2}\right)^{n/2}\exp\left(-\frac{n\mu^2}{2\sigma^2}\right)\exp\left[\underset{\sim}{w}'\left(\mu,\sigma^2\right)\underset{\sim}{T}(\underset{\sim}{x})\right]$$

şeklinde yazılabilir. Buradan, ortak olasılık yoğunluk fonksiyonu

$$h(\underset{\sim}{x})=\left(1/2\pi\right)^{n/2}\ \text{ve}\ g(T(\underset{\sim}{x});\mu,\sigma^2)=\left(\frac{1}{\sigma^2}\right)^{n/2}\exp\left(-\frac{n\mu^2}{2\sigma^2}\right)\exp\left[\underset{\sim}{w}'\left(\mu,\sigma^2\right)\underset{\sim}{T}(\underset{\sim}{x})\right]$$

fonksiyonları ile,

$$f(\underset{\sim}{x};\mu,\sigma^2)=g(\underset{\sim}{T}(\underset{\sim}{x});\mu,\sigma^2)\,h(\underset{\sim}{x})$$

şeklinde yazılır. Faktörizasyon teoremine göre, $\underset{\sim}{T}\left(\underset{\sim}{X}\right)=\left(\sum_{i=1}^{n}X_i,\ \sum_{i=1}^{n}X_i^2\right)$ şeklindeki iki boyutlu tahmin edici $(\mu,\sigma^2)$ için yeterlidir.

c) $X_1,X_2,\ldots,X_n$ parametresi $\theta$ olan düzgün dağılımdan bir örneklem olsun. $X$ lerin olasılık yoğunluk fonksiyonu,

$$f(x;\theta)=\begin{cases}1/\theta & , \quad 0<x<\theta \\ 0 & , \quad d.y.\end{cases}$$

şeklindedir. Bu olasılık yoğunluk fonksiyonu

$$I_A(x)=\begin{cases}1 & , \quad x\in A \\ 0 & , \quad x\notin A\end{cases}$$

gösterge fonksiyonu ile, $f(x;\theta)=\theta^{-1}I_{(0<x<\theta)}(x)$ şeklinde de yazılabilir. Buna göre, $X$ lerin ortak olasılık yoğunluk fonksiyonu $x_{(n)}=\max\{x_1,x_2,...,x_n\}$ için

$$\begin{aligned}f(\underset{\sim}{x};\theta)&=\prod_{i=1}^{n}f(x_i;\theta)=\frac{1}{\theta}I_{\{0<x_{(1)}<\theta\}}(x_1)\frac{1}{\theta}I_{\{0<x_{(2)}<\theta\}}(x_2)...\frac{1}{\theta}I_{\{0<x_{(n)}<\theta\}}(x_n)\\&=\frac{1}{\theta^n}I_{\{0<x_{(n)}<\theta\}}(\underset{\sim}{x})\end{aligned}$$

dir. Buradan $g(T(\underset{\sim}{x});\theta)=\theta^{-n}I_{\{0<x_{(n)}<\theta\}}(\underset{\sim}{x})$ ve $h(\underset{\sim}{x})=1$ fonksiyonları ile faktörizasyon teoreminden $T(\underset{\sim}{X})=X_{(n)}$ istatistiği $\theta$ için yeterlidir.

d) Olasılık yoğunluk fonksiyonu,

$$f(x;\theta)=\begin{cases}\theta x^{\theta-1} & , \quad 0<x<1 \\ 0 & , \quad d.y.\end{cases}$$

olan kitleden $n-$birimlik bir örneklem $X_1,X_2,\ldots,X_n$ olsun. Ortak olasılık yoğunluk fonksiyonu

$$f(\underset{\sim}{x};\theta)=\prod_{i=1}^{n}f(x_i;\theta)=\theta^n\prod_{i=1}^{n}x_i^{\theta-1}=\theta^n\left(\prod_{i=1}^{n}x_i\right)^{\theta}\prod_{i=1}^{n}\frac{1}{x_i}=g\left(T(\underset{\sim}{x});\theta\right)h(\underset{\sim}{x})$$

olarak yazılabildiğinden, $T(\underset{\sim}{X})=\prod_{i=1}^{n}X_i$ tahmin edicisi, $\theta$ için yeterlidir. Burada,

$$g\left(T(\underset{\sim}{x});\theta\right)=\theta^n\left(\prod_{i=1}^{n}x_i\right)^{\theta} \quad \text{ve} \quad h(\underset{\sim}{x})=\prod_{i=1}^{n}\frac{1}{x_i}$$

dir⊕

Faktörizasyon teoreminin uygulanabilmesi için $X_1,X_2,\ldots,X_n$ örnekleminin bağımsız aynı dağılımlı rasgele değişkenler olması zorunlu değildir. Ancak, daha önce örneklem denildiğinde bağımsız aynı dağılımlı rasgele değişkenlerin bir dizisini anlayacağımızı söylenmiştik. Teoremin ispatında da, rasgele değişkenlerin bağımsızlığı dikkate alınmamıştır.

Herhangi bir $\theta$ parametresi için bir tane yeterli istatistik bulunduğunda, bu yeterli istatistiğin her birebir fonksiyonu da yeterlidir. Bunun için, $T$ yeterli bir istatistik, $r$ de herhangi bir birebir fonksiyon olsun. Bu durumda, $T^*=r(T)$ de $\theta$ için yeterlidir. $T$ yeterli olduğundan öyle $g(t;\theta)$

ve $h(x)$ fonksiyonları vardır ki ortak olasılık yoğunluk fonksiyonu, $f(\underset{\sim}{x};\theta = g(T(\underset{\sim}{x});\theta)\,h(\underset{\sim}{x})$ şeklinde yazılabilir. Ayrıca $r$ birebir olduğundan, $T = r^{-1}(T^*)$ dir. Buradan, $f(\underset{\sim}{x};\theta)$ ortak olasılık veya olasılık yoğunluk fonksiyonu

$$\begin{aligned} f(\underset{\sim}{x};\theta) &= g(T(\underset{\sim}{x});\theta)\,h(\underset{\sim}{x} = g(r^{-1}(T^*(\underset{\sim}{x}));\theta)\ h(\underset{\sim}{x}) \\ &= (g \circ r^{-1})(T^*(\underset{\sim}{x});\theta)h(\underset{\sim}{x}) = g^*(T^*(\underset{\sim}{x});\theta)\,h(\underset{\sim}{x}) \end{aligned}$$

şeklinde de yazılabilir. Faktörizasyon teoremine göre, $T^* = r(T)$ de $\theta$ için yeterlidir. Buna göre, Örnek (7.1.2a) daki $T = \sum_{i=1}^{n} X_i$ istatistiği $p$ için yeterli olduğundan $\bar{X}_n$ de $p$ için yeterlidir. Benzer şekilde, (b) deki $\underset{\sim}{T}(\underset{\sim}{X}) = \left(\sum_{i=1}^{n} X_i,\ \sum_{i=1}^{n} X_i^2\right)$ tahmin edicisi $(\mu,\sigma^2)$ için yeterli olup $\underset{\sim}{T}^*(\underset{\sim}{X}) = (\bar{X}_n, S_n^2)$ iki boyutlu istatistiği de $(\mu,\sigma^2)$ için yeterlidir. Ayrıca, (d) de verilen $T = \prod_{i=1}^{n} X_i$ istatistik yeterli olduğundan, $T^* = -\sum_{i=1}^{n} \ln(X_i)$ tahmin edicisi de yeterlidir.

## 7.2. Minimal Yeterlilik

Bir önceki kısımda, bir parametre için yeterli istatistiğin tek olmadığını, yeterli istatistiğin her birebir fonksiyonunun da aynı parametre için yeterli olduğunu gördük.

**Tanım 7.2.1** (*minimal yeterli istatistik*) $T^*$ istatistiği herhangi bir $\theta$ parametresi için yeterli olsun. Eğer bir $T$ tahmin edicisi, $T^*$ yeterli istatistiğinin bir fonksiyonu olarak yazılabiliyorsa, $T$ ye $\theta$ için *minimal yeterlidir* denir⊗

Bir parametre için, herhangi bir yeterli istatistik bulunduğunda (faktörizasyon teoreminden), minimal yeterli istatistik bu istatistiğin bir fonksiyonudur. Tanımdan da anlaşıldığı gibi, bir parametre için minimal yeterli istatistiklerinin sınıfı, yeterli istatistiklerin sınıfını kapsar. Yani, herhangi bir parametre için yeterli istatistiklerden daha fazla minimal yeterli istatistik bulunabilir (Cacella ve Berger, 2002).

Faktörizasyon teoreminden yeterli istatistiği bulabiliyoruz. Minimal yeterli istatistiği bulmak için de aşağıdaki teorem kullanılabilir. Bu teoremin ispatı burada verilmemiştir (ispat için Casella ve Berger (2002) ye bakılabilir).

**Teorem 7.2.1** $X_1, X_2, \ldots, X_n$ olasılık veya olasılık yoğunluk fonksiyonu $f(x;\theta)$ olan kitleden bir örneklem olsun. Eğer

$$\frac{f(\underset{\sim}{x};\theta)}{f(\underset{\sim}{y};\theta)} = g(\underset{\sim}{x},\underset{\sim}{y})\ ,\ \theta \text{ dan bağımsız} \Leftrightarrow T(\underset{\sim}{x}) = T(\underset{\sim}{y})$$

önermesi iki yönlü olarak sağlanıyorsa, $T$ tahmin edicisi $\theta$ için minimal yeterlidir◊

**Örnek 7.2.1** $X_1, X_2, \ldots, X_n$ beklenen değeri $\mu$ varyansı $\sigma^2$ olan normal dağılımdan bir örneklem olsun. Kolayca görüleceği gibi,

$$\frac{f\left(\underset{\sim}{x};\mu,\sigma^2\right)}{f\left(\underset{\sim}{y};\mu,\sigma^2\right)}=\frac{\left(\frac{1}{2\pi\sigma^2}\right)^{n/2}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}x_i^2+\frac{\mu}{\sigma^2}\sum_{i=1}^{n}x_i-\frac{n\mu^2}{2\sigma^2}\right]}{\left(\frac{1}{2\pi\sigma^2}\right)^{n/2}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}y_i^2+\frac{\mu}{\sigma^2}\sum_{i=1}^{n}y_i-\frac{n\mu^2}{2\sigma^2}\right]}=\frac{\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}x_i^2+\frac{\mu}{\sigma^2}\sum_{i=1}^{n}x_i\right]}{\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}y_i^2+\frac{\mu}{\sigma^2}\sum_{i=1}^{n}y_i\right]}$$

$$=\exp\left(-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}x_i^2-\sum_{i=1}^{n}y_i^2\right]+\frac{\mu}{\sigma^2}\left[\sum_{i=1}^{n}x_i-\sum_{i=1}^{n}y_i\right]\right)$$

olup bu oranın parametrelerden ( $\mu$ ve $\sigma^2$ ) bağımsız olabilmesi için gerek ve yeter koşul,

$$\sum_{i=1}^{n}x_i^2=\sum_{i=1}^{n}y_i^2 \quad \text{ve} \quad \sum_{i=1}^{n}x_i=\sum_{i=1}^{n}y_i$$

olmasıdır. Yani, $\underset{\sim}{T}\left(\underset{\sim}{x}\right)=\left(\sum_{i=1}^{n}x_i, \sum_{i=1}^{n}x_i^2\right)$ olmak üzere, $\dfrac{f(\underset{\sim}{x};\mu,\sigma^2)}{f(\underset{\sim}{y};\mu,\sigma^2)}$ oranı $\mu$ ve $\sigma^2$ den bağımsız olması için gerek ve yeter koşul $\underset{\sim}{T}(\underset{\sim}{x})=\underset{\sim}{T}(\underset{\sim}{y})$ olmasıdır. Teorem (7.2.1) e göre, $\underset{\sim}{T}\left(\underset{\sim}{X}\right)=\left(\sum_{i=1}^{n}X_i, \sum_{i=1}^{n}X_i^2\right)$ istatistiği $(\mu,\sigma^2)$ için minimal yeterlidir. Ayrıca,

$$\underset{\sim}{T}^*\left(\underset{\sim}{X}\right)=\left(\frac{1}{n}\sum_{i=1}^{n}X_i, \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2\right)=\left(\bar{X}_n,\, S_n^2\right)$$

de $(\mu,\sigma^2)$ için minimal yeterlidir.

## 7.3. Yardımcı İstatistik (Ancillary Statistic)

**Tanım 7.3.1** Dağılımı parametreye bağlı olmayan istatistiğe *yardımcı istatistik* (ancillary statistic) denir⊗

Tanımdan da anlaşıldığı gibi, yardımcı istatistiğinin dağılımı parametreye bağlı olmadığından, yardımcı istatistiğe bağlı olarak parametre hakkında herhangi bir sonuç çıkarım yapılamaz. Ancak yardımcı istatistik başka bazı özellikler ile beraber kullanıldığında parametre hakkında önemli bilgiler elde edilir.

**Örnek 7.3.1** Olasılık yoğunluk fonksiyonu $f(x;\theta)$ olan düzgün dağılımdan bir örneklem $X_1, X_2, \ldots, X_n$ olsun. Burada,

$$f(x;\theta)=\begin{cases}1 & , \quad \theta<x<\theta+1\\ 0 & , \quad d.y.\end{cases}$$

olmak üzere, $X$ lerin dağılım fonksiyonu,

$$F(x;\theta)=\begin{cases}0 & , \quad x\le\theta\\ x-\theta & , \quad \theta<x<\theta+1\\ 1 & , \quad x\ge\theta+1\end{cases}$$

şeklindedir. $\mathcal{R}=X_{(n)}-X_{(1)}$ örneklem uzunluğu istatistiğinii gözönüne alalım. $\mathcal{R}$ nin olasılık yoğunluk fonksiyonu için $X_{(1)}$ ve $X_{(n)}$ sıra istatistiklerinin ortak olasılık yoğunluk fonksiyonu,

$$g(x,y)=\begin{cases}n(n-1)(y-x)^{n-2} & , \quad \theta<x<y<\theta+1\\ 0 & , \quad d.y.\end{cases}$$

olup $\mathcal{R}=X_{(n)}-X_{(1)}$ nin olasılık yoğunluk fonksiyonu için $M=(X_{(1)}+X_{(n)})/2$ yardımcı dönüşümünü tanımlayalım. Ters dönüşümler, $X_{(1)}=(2M-\mathcal{R})/2$ ve $X_{(n)}=(2M+\mathcal{R})/2$ olup Jacobien matrisinin determinantı 1 dir. Dolayısı ile, $\mathcal{R}$ ve $M$ nin ortak olasılık yoğunluk fonksiyonu,

$$f_{\mathcal{R}\mathcal{M}}(r,m)=\begin{cases}n(n-1)r^{n-2} & , \quad 0<r<1\,,\ \theta+\frac{r}{2}<m<\theta+1-\frac{r}{2}\\ 0 & , \quad d.y.\end{cases}$$

olarak elde edilir. Ortak olasılık yoğunluk fonksiyonunun $D_M$ üzerinden integrali

$$\int_{m=\theta+r/2}^{\theta+1-r/2} n(n-1)r^{n-2}dm=n(n-1)r^{n-2}\left(\left[\theta+1-r/2\right]-\left[\theta+r/2\right]\right)=n(n-1)r^{n-2}(1-r)$$

olup $\mathcal{R}$ nin olasılık yoğunluk fonksiyonu,

$$f_{\mathcal{R}}(r)=\begin{cases}n(n-1)r^{n-2}(1-r) & , \quad 0<r<1\\ 0 & , \quad d.y.\end{cases}$$

olarak bulunur. Görüldüğü gibi $\mathcal{R}=X_{(n)}-X_{(1)}$ nin dağılımı $\theta$ parametresine bağlı değildir. Yani, $\mathcal{R}$ istatistiği $\theta$ için bir yardımcı istatistiktir. Benzer şekilde, $X_1,X_2,\ldots,X_n$ $N(\mu,\sigma^2)$ dağılımından bir örneklem ise, $(n-1)S_n^2/\sigma^2\sim\chi_{n-1}^2$ olup $S_n^2$ örneklem varyansının dağılımı $\mu$ ye bağlı değildir. Yani, $S_n^2$ istatistiği $\mu$ için bir yardımcı istatistiktir$\oplus$

### 7.4. Tamlık (Completeness)

Örnek (7.3.1) de $X_1,X_2,\ldots,X_n\sim U(\theta,\theta+1)$ örneklemi için $\mathcal{R}=X_{(n)}-X_{(1)}$ örneklem uzunluğu istatistiğinin bir yardımcı istatistik olduğunu gördük. Ayrıca, $(\mathcal{R},M)$ de $\theta$ için minimal yeterli olup yardımcı istatistik, minimal yeterli istatistiğin önemli bir bileşenidir. Genellikle, yeterli istatistik ile yardımcı istatistik bağımsız değildir. Ancak, minimal yeterli istatistik aynı zamanda tam ise minimal yeterli istatistik ile yardımcı istatistik bağımsızdır (Teorem (7.4.2)).

**Tanım 7.4.1** $X_1, X_2, \ldots, X_n$ olasılık veya olasılık yoğunluk fonksiyonu $f(x;\theta)$ olan kitleden bir örneklem, $T$ de $\theta$ için yeterli olsun. Bütün $\theta$ lar için

$$E_\theta(g(T)) = 0 \Rightarrow P_\theta(g(T)=0) = 1$$

önermesi sağlanıyorsa, $T$ nin olasılık yoğunluk fonksiyonlarının ailesi *tamdır* denir $\otimes$

Buradaki, olasılık yoğunluk fonksiyonlarının ailesinin tamlığı yerine kısaca $T$ tamdır diyeceğiz. Ayrıca tamlık, olasılık dağılımlarının ailesi için bir kavram olup özel dağılımlar için geçerli değildir. Örneğin, $X \sim N(0,1)$ ise, $g(x) = x$ için $E(g(X)) = 0$ olmasına rağmen $P(g(X)=0) = 0$ dir. Ancak, $X \sim N(\theta,1)$ olsaydı, bütün $\theta$ lar için $E_\theta(g(X)) = 0$ olacak şekilde bir $g$ fonksiyonu bulunamaz. Yani, $E_\theta\left(g(X)\right) = 0 \Rightarrow P_\theta\left(g(X)=0\right) = 1$ önermesi sağlanır. Dolayısı ile, $X \sim N(\theta,1)$ dağılımlar ailesi tamdır.

**Örnek 7.4.1** a) $X_1, X_2, \ldots, X_n \sim Bern(p)$ olsun. $T = X_1 + X_2 + \ldots + X_n$ tahmin edicisi $p$ için yeterli ve $T \sim Binom(n,p)$ olduğunu biliyoruz. Şimdi, Binom dağılımlar ailesinin (yani $T$ nin) tam olduğunu gösterelim. Bunun için, bütün $p$ ler için $E_p(g(T)) = 0$ olsun. Buradan,

$$0 = E_p\left(g(T)\right) = \sum_{t=0}^{n} g(t) P_p(T=t) = \sum_{t=0}^{n} g(t)\binom{n}{t} p^t (1-p)^{n-t} = (1-p)^n \sum_{t=0}^{n} g(t)\binom{n}{t}\left(\frac{p}{1-p}\right)^t \text{ ise}$$

bütün $p$ ler için $(0 < p < 1)$

$$\sum_{t=0}^{n} g(t)\binom{n}{t}\left(\frac{p}{1-p}\right)^t = \sum_{t=0}^{n} g(t)\binom{n}{t} r^t = 0$$

olduğundan

$$E\left(g(T)\right) = 0 \text{ ise } \sum_{t=0}^{n} g(t)\binom{n}{t} r^t = 0$$

dır. $r = p/(1-p) > 0$ ve toplam içindeki bütün terimler pozitif olduğundan toplamın sıfır olması için bütün $t$ ler için $g(t) = 0$ olması gerekir. Yani,

$$E_p(g(T)=0) = 0 \Rightarrow P_p(g(T)=0) = 1$$

önermesi doğrudur. Böylece, $T$ istatistiği veya Bernoulli dağılımlar ailesinin tam olduğu söylenebilir.

b) $X_1, X_2, \ldots, X_n$ parametresi $\theta$ olan Poisson dağılımından bir örneklem olsun. $T = \sum_{i=1}^{n} X_i$ istatistiği $\theta$ için yeterli olup $T \sim Poisson(n\theta)$ dır. $T$ nin olasılık fonksiyonu, $t = 0,1,2,3,\ldots$ için $g(t,\theta) = P_\theta(T=t) = e^{-n\theta}(n\theta)^t / t!$ şeklindedir. Bu şekildeki olasılık fonksiyonlarının ailesi $\mathcal{C} = \{g(t,\theta) : \theta > 0\}$ olsun. $E_\theta(h(T)) = 0$ olacak şekilde bir $h$ fonksiyonu belirleyelim. $T$ nin tam olduğunu göstermek için her $t$ için $h(t) = 0$ olduğunun gösterilmesi gerekir. Dolayısı ile,

$$0 = E_\theta\left(h(T)\right) = \sum_{t=0}^{\infty} h(t) e^{-n\theta}(n\theta)^t / t!$$

$$= e^{-n\theta}\left[h(0) + h(1)\frac{n\theta}{1} + h(2)\frac{(n\theta)^2}{2!} + h(3)\frac{(n\theta)^3}{3!} + \ldots\right]$$

olduğundan,

$$\left[h(0) + \left[n\,h(1)\right]\theta + \left[\frac{n^2}{2!}h(2)\right]\theta^2 + \left[\frac{n^3}{3!}h(3)\right]\theta^3 + \ldots\right] = 0$$

yazılır. Bütün pozitif $\theta$ lar ($\theta > 0$) için toplamı sıfır olan herhangi bir serinin bütün katsayıları sıfırdır (Hogg ve Craig, 1995). Böylece, bütün pozitif $\theta$ lar için $h(t) = 0, \forall t$ yazılır. Buradan, $P_\theta(h(T)=0) = 1$ olup $E_\theta(h(T)=0) = 0 \Rightarrow P_\theta(h(T)=0) = 1$ önermesi sağlanır. Yani, Poisson dağılımlar ailesi tamdır $\oplus$

Bir dağılımlar ailesinin veya istatistiğin tam olduğunun gösterilmesi bazen kolay olmayabilir. Bu dağılımlar ailesi aşağıda tanımı verilen bir üstel aile ise tahmin edicinin tam olduğu kolayca görülebilir. Üstel aileler beşinci bölümde kısaca incelenmişti. Tanım (5.4.1) deki üstel aile tanımına benzer bir tanım aşağıda verilmiştir.

**Tanım 7.4.2** Olasılık veya olasılık yoğunluk fonksiyonlarının $\mathcal{G} = \{f(x;\theta) : \theta \in \Theta\}$ şeklindeki bir ailesi, her $f(x;\theta) \in \mathcal{G}$ için

$$f(x;\theta) = h(x)\,c(\theta)\exp(t(x)\,w(\theta))$$

şeklinde yazılabiliyorsa, $\mathcal{G} = \{f(x;\theta) : \theta \in \Theta\}$ ailesine bir *üstel aile* denir $\otimes$

**Teorem 7.4.1** $X_1, X_2, \ldots, X_n$ üstel aile özelliğine sahip kitleden bir örneklem olsun. $X$ lerin ortak olasılık veya olasılık yoğunluk fonksiyonu $f(\underset{\sim}{x};\theta)$,

$$f(\underset{\sim}{x};\theta) = h(\underset{\sim}{x})\,c(\theta)\exp\left(t(\underset{\sim}{x})\,w(\theta)\right)$$

özelliğine sahip ise $T = \sum_{i=1}^{n} t(X_i)$ yeterli ve tamdır (Hogg ve Craig, 1995, sayfa 335) $\Diamond$

**Örnek 7.4.2** a) Bir önceki örnekte Bernoulli dağılımının parametresi için önerilen $T$ nin tam olduğunu gördük. $T \sim Binom(n, p)$ olup olasılık fonksiyonu,

$$f(t;p) = P_p(T=t) = \binom{n}{t} p^t (1-p)^{n-t} = \binom{n}{t}(1-p)^n \left(\frac{p}{1-p}\right)^t = \binom{n}{t}(1-p)^n \exp\left(t \ln\left(\frac{p}{1-p}\right)\right)$$

olarak yazılabilir. Yani, $f(t;p)$ Binom olasılık fonksiyonlarının ailesi üsteldir. O halde Teorem (7.4.1) gereğince $T$ yeterli ve tamdır.

Benzer şekilde Poisson dağılımlar ailesi de üstel olup tamdır. $T \sim Poisson(n\theta)$ olup $T$ nin olasılık fonksiyonu

$$g(t,\theta) = P_\theta(T=t) = e^{-n\theta}(n\theta)^t / t! = n^t e^{-n\theta} \exp(\ln(\theta^t)) / t! = n^t e^{-n\theta} \exp(t \ln(\theta)) / t!$$

şeklinde yazılabilir. Böylece, $T = \sum_{i=1}^{n} X_i$ nin $\theta$ için yeterli ve tam olduğu söylenebilir.

$X_1, X_2, \ldots, X_n \sim Üstel(\theta)$ olsun. $X$ lerin olasılık yoğunluk fonksiyonu,

$$f(x;\theta) = \begin{cases} \dfrac{1}{\theta} e^{-x/\theta} & , \quad x > 0 \\ 0 & , \quad d.y. \end{cases}$$

olup $f(x;\theta) = \theta^{-1} \exp(-x\theta)$ olduğundan üstel aile formunda yazılabilir. Dolayısı ile, $t(x) = x$ olmak üzere, $T = \sum_{i=1}^{n} X_i$ tahmin edicisi $\theta$ için yeterli ve tamdır.

b) Bazı olasılık yoğunluk fonksiyonları üstel aile formunda yazılamaz. Örneğin, $X_1, X_2, \ldots, X_n \sim U(0, \theta)$ olsun. $X$ lerin olasılık yoğunluk fonksiyonu,

$$f(x;\theta) = \theta^{-1} I_{(0<x<\theta)}(x)$$

şeklinde olup, üstel aile formunda yazılamaz. Daha önce gösterildiği gibi, $X_{(n)}$ sıra istatistiği $\theta$ için yeterlidir (Örnek (7.1.2c)). Şimdi tam olduğunu gösterelim. Bunun için $T = X_{(n)}$ nin olasılık yoğunluk fonksiyonu $0 < t < \theta$ için $f(t;\theta) = n\theta^{-n} t^{n-1}$ olup, $E_\theta(g(T)) = 0$ olacak şekilde bir $g$ fonksiyonu belirleyelim. $E_\theta(g(T)) = 0$ ise $\frac{d}{d\theta} E_\theta(g(T)) = 0$ olup çarpımın türevi kuralından $\frac{d}{d\theta} E_\theta(g(T)) = 0$ ifadesi

$$0 = E_\theta\left(g(T)\right) = \frac{d}{d\theta}E_\theta\left(g(T)\right) = \frac{d}{d\theta}\int_0^\theta g(t)\, f(t;\theta)\,dt = \frac{d}{d\theta}\int_0^\theta g(t)\, n\,\theta^{-n}\, t^{n-1}\,dt$$

$$= n\,\theta^{-n}\,\frac{d}{d\theta}\int_0^\theta g(t)\; t^{n-1}\,dt + \int_0^\theta g(t)\; t^{n-1}\,dt\left[\frac{d}{d\theta}\left(n\,\theta^{-n}\right)\right]$$

olarak yazılabilir. Diğer taraftan, $E_\theta(g(T)) = 0$ olduğundan $\int_0^\theta g(t)\; t^{n-1}\,dt = 0$ dır. Buradan da, son eşitliğin ikinci kısmı sıfır olup her $\theta$ için

$$0 = \frac{d}{d\theta}E_\theta\left(g(T)\right) = n\,\theta^{-n}\,\frac{d}{d\theta}\int_0^\theta g(t)\; t^{n-1}\,dt = n\,\theta^{-n}\; g(\theta)\,\theta^{n-1} = n\,\theta^{-1} g(\theta)$$

elde edilir. Yani, bütün pozitif $\theta$ ($\theta > 0$) lar için $E_\theta(g(T)) = 0$ ise $g(\theta) = 0$ dır. Buradan da, $P_\theta(g(T) = 0) = 1$ elde edilir. Yani, $T = X_{(n)}$ sıra istatistiği düzgün dağılımın $\theta$ parametresi için yeterli ve tamdır.

Benzer şekilde, $X_1, X_2, \ldots, X_n$ olasılık yoğunluk fonksiyonu $f(x;\theta)$ olan kitleden bir örneklem olsun. $X$ lerin olasılık yoğunluk fonksiyonu,

$$f\left(x;\theta\right) = \begin{cases} e^{-(x-\theta)} & , \quad x > \theta \\ \quad 0 & , \quad d.y \end{cases}$$

üstel aile formunda yazılamaz. Bu olasılık yoğunluk fonksiyonu $I_A(x)$ gösterge fonksiyonu ile $f\left(x;\theta\right) = e^{-(x-\theta)}\, I_{(x>\theta)}(x)$ şeklinde yazılabilir. Faktörizasyon teoreminden $T = X_{(1)}$ sıra istatistiği $\theta$ için yeterlidir.

Şimdi $X_{(1)}$ in tamlığını inceleyelim. Önce, $T$ nin olasılık yoğunluk fonksiyonunun,

$$f_T\left(t;\theta\right) = \begin{cases} n\; e^{-n\,(t-\theta)} & , \quad t > \theta \\ \quad 0 & , \quad d.y \end{cases}$$

şeklinde olduğunu belirtelim. $E_\theta(h(T)) = 0$ olacak şekilde bir $h$ fonksiyonu belirlensin. Buradan, $E_\theta(h(T)) = 0$ ise $E_\theta(h(T))$ nin $\theta$ ya göre birinci türevi de sıfırdır. $E_\theta(h(T)) = 0$ olduğundan $\int_\theta^\infty h(t)\, e^{-n\,t} dt = 0$ olup

$$0 = \frac{d}{d\theta}E_\theta\left(h(T)\right) = \frac{d}{d\theta}\int_\theta^\infty h(t)\, n\, e^{-n(t-\theta)}\,dt = n\,\frac{d}{d\theta}\left[e^{n\theta}\int_\theta^\infty h(t)\, e^{-nt} dt\right]$$

$$= n\left[\frac{1}{n}e^{n\theta}\int_\theta^\infty h(t)\, e^{-n\,t} dt - e^{n\theta}\; h(\theta)\, e^{-n\theta}\right] = -\frac{1}{n}h(\theta)$$

bulunur. Buna göre, her $\theta$ için $h(\theta)=0$ elde edilir. Yani, $P_\theta(h(T)=0)=1$ olup $T=X_{(1)}$ tamdır $\oplus$

Daha önce de bahsedildiği gibi, yardımcı istatistik tek başına parametre hakkında herhangi bir bilgi içermez. Ancak bu yardımcı istatistik, yeterli ve tam bir tahmin edici ile beraber kullanıldığında önemli istatistiki sonuçlara ulaşılır. Aşağıda ispatsız olarak verilen teorem bu söylenenleri özetlemektedir (ispat için Casella ve Berger (2002), sayfa 287 ye bakılabilir).

**Teorem 7.4.2** *(Basu Teoremi)* $T$ minimal yeterli ve tam ise, $T$ herhangi bir yardımcı istatistikten bağımsızdır $\lozenge$

**Örnek 7.4.3** $X_1, X_2, \ldots, X_n$ beklenen değeri $\theta$ olan üstel dağılımdan bir örneklem olsun. Bu durumda, $T = X_1 + X_2 + \ldots + X_n \sim Gamma(n, \theta)$ olduğunu biliyoruz. Ayrıca, $T$ yeterli ve tamdır (üstel aile özelliğini sağlar). $Y = X_1 + \ldots + X_{n-1} \sim Gamma(n-1,\ \theta)$ olup

$$U = \frac{X_n}{X_1 + X_2 + \ldots + X_n} = \frac{X_n}{Y + X_n} \sim Beta(1, n-1)$$

olduğu da gösterilebilir. Yani, $U$ yardımcı istatistik olup Basu teoremine göre $U$ ile $T$ bağımsızdır. Buna göre $U$ nun beklenen değeri

$$\theta = E_\theta(X_n) = E_\theta(UT) = E_\theta(U)E_\theta(T) = (\theta n)\, E_\theta(U) \Rightarrow E_\theta(U) = n^{-1}$$

şeklinde kolayca hesaplanabilir.

Diğer taraftan, $X_1, X_2, \ldots, X_n$ beklenen değeri $\mu$, varyansı $\sigma^2$ olan normal dağılımdan bir örneklem ise, $\bar{X}_n$ örneklem ortalaması $\mu$ için yeterli ve tamdır (üstel aile özelliğini sağlar). Ayrıca, $(n-1)S_n^2/\sigma^2 \sim \chi^2_{n-1}$ olduğunu (Teorem (6.3.1c) biliyoruz. Dolayısı ile $S_n^2$ nin dağılımı $\mu$ parametresine bağlı değildir. Yani, $S_n^2$ istatistiği $\mu$ için yardımcı istatistiktir. Basu teoremine göre de $\bar{X}_n$ ile $S_n^2$ istatistikleri bağımsızdır. Daha önce, Teorem (6.3.1b) de $\bar{X}_n$ ile $S_n^2$ nin bağımsız olduğu gösterildi. Oradaki ispat ile buradaki sonucu karşılaştırınız $\oplus$