EN ÇOK OLABİLİRLİK KESTİRİMİ

(y_i, x_i), $i = 1, 2, ..., n$ verileri ele alınsın. Eğer regresyon modelindeki hataların $NID(0, \sigma^2)$ olduğu varsayılırsa bu örneklemdeki y_i gözlemleri, $\beta_0 + \beta_1 x_i$ ortalamasına ve σ^2 varyansına sahip normal ve bağımsız dağılımlı raslantı değişkenleridir.

Normal hatalı basit doğrusal regresyon için olabilirlik fonksiyonu,

$$
\begin{aligned}
L(y_i, x_i, \beta_0, \beta_i, \sigma^2) &= \prod_{i=1}^{n} (2\pi\sigma^2)^{-1/2} \exp\left[-\frac{1}{2\sigma^2} (y_i - \beta_0 - \beta_1 x_i)^2 \right] \\
&= (2\pi\sigma^2)^{-1/2} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 \right] \qquad \textbf{(1.52)}
\end{aligned}
$$

En çok olabilirlik kestiricileri, L 'yi ya da eş değer olarak $\ln L$ 'yi maksimum yapan $\tilde{\beta}_0$, $\tilde{\beta}_1$ ve $\tilde{\sigma}^2$ gibi parametre değerleridir.

$$
\begin{aligned}
\ln L(y_i, x_i, \beta_0, \beta_i, \sigma^2) &= -\left(\frac{n}{2}\right) \ln 2\pi - \left(\frac{n}{2}\right) \ln \sigma^2 \\
&= -\left(\frac{1}{2\sigma^2}\right) \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 \qquad \textbf{(1.53)}
\end{aligned}
$$

olup $\tilde{\beta}_0$, $\tilde{\beta}_1$ ve $\tilde{\sigma}^2$ en çok olabilirlik kestiricileri aşağıdaki denklemleri sağlamalıdırlar:

$$
\left. \frac{\partial \ln L}{\partial \beta_0} \right|_{\tilde{\beta}_0, \tilde{\beta}_1, \tilde{\sigma}^2} = \frac{1}{\tilde{\sigma}^2} \sum_{i=1}^{n} (y_i - \tilde{\beta}_0 - \tilde{\beta}_1 x_i) = 0 \qquad \textbf{(1.54a)}
$$

$$
\left. \frac{\partial \ln L}{\partial \beta_1} \right|_{\tilde{\beta}_0, \tilde{\beta}_1, \tilde{\sigma}^2} = \frac{1}{\tilde{\sigma}^2} \sum_{i=1}^{n} (y_i - \tilde{\beta}_0 - \tilde{\beta}_1 x_i) x_i = 0 \qquad \textbf{(1.54b)}
$$

$$
\left. \frac{\partial \ln L}{\partial \sigma^2} \right|_{\tilde{\beta}_0, \tilde{\beta}_1, \tilde{\sigma}^2} = -\frac{n}{2\tilde{\sigma}^2} + \frac{1}{2\tilde{\sigma}^4} \sum_{i=1}^{n} (y_i - \tilde{\beta}_0 - \tilde{\beta}_1 x_i)^2 = 0 \qquad \textbf{(1.54c)}
$$

En çok olabilirlik kestiricileri,

$$\tilde{\beta}_0 = \overline{y} - \tilde{\beta}_1 \overline{x} \qquad \textbf{(1.55a)}$$

$$\tilde{\beta}_1 = \frac{\sum_{i=1}^{n} y_i (x_i - \overline{x})}{\sum_{i=1}^{n} (x_i - \overline{x})^2} \qquad \textbf{(1.55b)}$$

$$\tilde{\sigma}^2 = \frac{\sum_{i=1}^{n} (y_i - \tilde{\beta}_0 - \tilde{\beta}_1 x_i)^2}{n} \qquad \textbf{(1.55c)}$$

olarak elde edilmiştir. Kesim noktasının ve eğimin en çok olabilirlik kestiricileri olan $\tilde{\beta}_0$ ve $\tilde{\beta}_1$, bu parametrelerin en küçük kareler kestiricileriyle özdeştir. $\tilde{\sigma}^2$, σ^2'nin yanlı bir kestiricisi olup $\hat{\sigma}^2$ ile $\tilde{\sigma}^2 = \left[(n-1)/n\right]\hat{\sigma}^2$ şeklinde bir bağlantısı vardır.

En çok olabilirlik kestiricileri, en küçük kareler kestiricilerinden daha iyi istatistiksel özelliklere sahiptir.

- En çok olabilirlik kestiricileri **yansızdır** ve diğer bütün yansız kestiricilerle karşılaştırıldığında **minimum varyansa** sahiptir. Ayrıca **tutarlı** kestiricilerdir ve **yeterli** istatistiklerdir.

En çok olabilirlik kestirimi, en küçük kareler kestiricilerinden daha kesin istatistiksel varsayımlar gerektirmektedir.

- En küçük kareler kestiricileri, yalnızca ikinci moment varsayımlarını (beklenen değer, rasgele hatalar arasındaki varyanslar ve kovaryanslara ilişkin varsayımlar) gerektirirken, en çok olabilirlik kestiricileri ise tam bir dağılım varsayımı gerektirmektedir.

BAĞIMSIZ DEĞİŞKENİN RASTGELE OLDUĞU DURUM

Bileşik Dağılımlı x ve y

x ve y'nin bileşik dağılıma sahip raslantı değişkenleri olduğu ancak bu bileşik dağılımın bilinmediği varsayılsın.

1) x verildiğinde y'nin koşullu dağılımı, $\beta_0 + \beta_1 x$ koşullu ortalaması ve σ^2 koşullu varyansı ile normaldir.
2) x'ler, olasılık dağılımının β_0, β_1 ya da σ^2'yi içermediği bağımsız raslantı değişkenleridir.

Bu koşullar sağlandığı durumda bütün regresyon işlemleri değişmeden kalırken güven katsayıları ve istatistiksel hatalar farklı bir yorum kazanır.

Bileşik Normal Dağılımlı x ve y

y ve x'in **iki değişkenli normal dağılıma** göre bileşik dağıldığı varsayılsın:

$$f(y,x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{y-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{y-\mu_1}{\sigma_1}\right)\left(\frac{x-\mu_2}{\sigma_2}\right)\right]\right\} \tag{1.56}$$

Burada μ_1 ve σ_1^2, y'nin ortalama ve varyansıdır; μ_2 ve σ_2^2, x'in ortalama ve varyansıdır ve

$$\rho = \frac{E(y-\mu_1)(x-\mu_2)}{\sigma_1\sigma_2} = \frac{\sigma_{12}}{\sigma_1\sigma_2}$$

değeri y ile x arasındaki **korelasyon katsayısıdır**. σ_{12} terimi y ve x'in **kovaryansıdır.**

x'in verilen bir değeri için y'nin **koşullu dağılımı** şöyledir:

$$f(y/x) = \frac{1}{\sqrt{2\pi}\sigma_{1.2}} \exp\left[-\frac{1}{2}\left(\frac{y-\beta_0-\beta_1 x}{\sigma_{1.2}}\right)^2\right] \tag{1.57}$$

Buradan,

$$\beta_0 = \mu_1 - \mu_2\rho\frac{\sigma_1}{\sigma_2} \tag{1.58a}$$

$$\beta_1 = \frac{\sigma_1}{\sigma_2}\rho \tag{1.58b}$$

$$\sigma^2{}_{1.2} = \sigma_1^2(1-\rho^2) \tag{1.58c}$$

olarak elde edilir. x verildiğinde y'nin koşullu dağılımı,

$$E(y/x) = \beta_0 + \beta_1 x \tag{1.59}$$

koşullu ortalaması ve $\sigma^2{}_{1.2}$ koşullu varyansıyla normaldir.

ρ korelasyon katsayısı ve β_1 arasında bir ilişki vardır. $\rho = 0$ olduğunda $\beta_1 = 0$ olur ve bu da y 'nin x üzerinde doğrusal bir regresyonu olmadığı anlamına gelmektedir.

β_0 ve β_1 parametrelerinin kestiriminde en çok olabilirlik yöntemi kullanılabilir. Bu parametrelerin en çok olabilirlik kestiricileri,

$$\hat{\beta}_0 = \overline{y} - \hat{\beta}_1 \overline{x} \qquad \textbf{(1.60a)}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i (x_i - \overline{x})}{\sum_{i=1}^{n} (x_i - \overline{x})^2} = \frac{S_{xy}}{S_{xx}} \qquad \textbf{(1.60b)}$$

olup bu kestiriciler, x'in kontrol edilebilir bir değişken olarak kabul edildiği en küçük kareler yöntemindeki kestiricilerle özdeştir.

- Genel olarak, y ve x'in bileşik normal dağıldığı bir regresyon modeli, x'in kontrol edilebilir bir değişken olarak kabul edildiği önceki model için kullanılan yöntemlerle analiz edilebilir. Bunun nedeni, x verildiğinde y raslantı değişkeninin $\beta_0 + \beta_1 x$ ortalamasıyla ve $\sigma^2{}_{1.2}$ varyansıyla bağımsız ve normal dağılmasıdır.

ρ 'nun kestiricisi **örneklem korelasyon katsayısıdır:**

$$r = \frac{\sum_{i=1}^{n} y_i (x_i - \overline{x})}{\left[\sum_{i=1}^{n} (x_i - \overline{x})^2 \sum_{i=1}^{n} (y_i - \overline{y})^2 \right]^{1/2}} = \frac{S_{xy}}{\left[S_{xx} SS_T \right]^{1/2}} \qquad \textbf{(1.61)}$$

Aşağıdaki denklem göz önüne alınsın:

$$\hat{\beta}_1 = \left(\frac{SS_T}{S_{xx}} \right)^{1/2} r \qquad \textbf{(1.62)}$$

$\hat{\beta}_1$ ve r, farklı bilgiler vermesine rağmen yakından ilişkilidir. r, örneklem korelasyon katsayısı x ve y arasındaki doğrusal ilişkinin bir ölçüsüyken $\hat{\beta}_1$ ise x'teki bir birimlik değişimde y'nin ortalamasındaki değişikliği ölçmektedir.

$$r^2 = \hat{\beta}_1^2 \frac{S_{xx}}{SS_T} = \frac{\hat{\beta}_1 S_{xy}}{SS_T} = \frac{SS_R}{SS_T} = R^2$$

olmak üzere görüldüğü gibi belirtme katsayısı R^2, y ve x arasındaki korelasyonun karesidir.

Korelasyon katsayısının sıfıra eşit olup olmadığı,

$$H_0 : \rho = 0 \ , \ H_1 : \rho \neq 0 \tag{1.63}$$

hipotezi test edilerek belirlenir. Bu test için uygun test istatistiği,

$$t_0 = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \tag{1.64}$$

olup bu test istatistiği, $H_0 : \rho = 0$ gerçekte doğru iken (n-2) serbestlik dereceli t dağılımlıdır. $|t_0| > t_{\alpha/2,n-2}$ ise sıfır hipotezi reddedilecektir. Bu test, $H_0 : \beta_1 = 0$ için verilen teste eş değerdir. Bu eş değerlik Denklem (1.62)'den kaynaklanmaktadır.

$$H_0 : \rho = \rho_0 \quad , \quad H_1 : \rho \neq \rho_0 \tag{1.65}$$

$\rho_0 \neq 0$ olması durumunda daha büyük örneklemler için ($n \geq 25$) Z istatistiği,

$$Z = \operatorname{arctan} h\, r = \frac{1}{2} \ln \frac{1+r}{1-r} \tag{1.66}$$

olup $\mu_Z = \operatorname{arctan} h\, \rho = \frac{1}{2} \ln \frac{1+\rho}{1-\rho}$ ortalaması ve $\sigma_Z{}^2 = (n-3)^{-1}$ varyansı ile yaklaşık normal dağılımlıdır.

$$Z_0 = (\operatorname{arctan} h\, r - \operatorname{arctan} h\, \rho_0)(n-3)^{1/2} \tag{1.67}$$

olmak üzere, $|Z_0| > Z_{\alpha/2}$ ise $H_0 : \rho = \rho_0$ hipotezi reddedilir.

ρ katsayısı için yüzde 100(1-α) güven aralığı,

$$\tanh\left(\operatorname{arctan} h\, r - \frac{Z_{\alpha/2}}{\sqrt{n-3}} \right) \leq \rho \leq \tanh\left(\operatorname{arctan} h\, r - \frac{Z_{\alpha/2}}{\sqrt{n-3}} \right) \tag{1.68}$$

olup burada $\tanh u = (e^u - e^{-u}) / (e^u + e^{-u})$ ile de tanımlıdır.

Örnek 1.9 Teslim Süresi Verileri

TABLO 1.8 Örnek 1.9 Verileri

Gözlem	Teslim Süresi, y	Teslim Hacmi, x	Gözlem	Teslim Süresi, y	Teslim Hacmi, x
1	16.68	7	14	19.75	6
2	11.50	3	15	24.00	9
3	12.03	3	16	29.00	10
4	14.88	4	17	15.35	6
5	13.75	6	18	19.00	7
6	18.11	7	19	9.50	3
7	8.00	2	20	35.10	17
8	17.83	7	21	17.90	10
9	79.24	30	22	52.32	26
10	21.50	5	23	18.75	9
11	40.33	16	24	19.83	8
12	21.00	10	25	10.75	4
13	13.50	4			

Teslim süresi y ve teslim hacmi x arasındaki örneklem korelasyon katsayısı,

$$r = \frac{S_{xy}}{\left[S_{xx} SS_T\right]^{1/2}} = \frac{2473.3440}{\left[(1136.5600)(5784.5426)\right]^{1/2}} = 0.9646$$

TABLO 1.9 Meşrubat Teslim Süresi Verileri İçin Minitab Çıktısı

Regression Analysis : Time versus Cases

The regression equation is
Time = 3.32 + 2.18 Cases

Predictor	Coef	SE Coef	T	P
Constant	3.321	1.371	2.42	0.024
Cases	2.176	0.124	17.55	0.000

S = 4.18140 R - Sq = 93.0% R-Sq(adj) = 92.7%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	5382.4	5382.4	307.85	0.000

Residual Error	23	402.1	17.5
Total	24	5784.5	

Eğer teslim süresi ve teslim hacminin bileşik normal dağıldığı varsayıldığında,

$$H_0 : \rho = 0 \ , \quad H_1 : \rho \neq 0$$

hipotezleri test edilir ve test istatistiği,

$$t_0 = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.9646\sqrt{23}}{\sqrt{1-0.9305}} = 17.55$$

olup $|17.55| > t_{0.025,23} = 2.069$ olduğundan H_0 hipotezi reddedilir.

$\operatorname{arctan} h\, r = \operatorname{arctan} h\, 0.9646 = 2.0082$ olmak üzere ρ için yaklaşık % 95 güven aralığı,

$$\tanh\left(2.0082 - \frac{1.96}{\sqrt{22}}\right) \leq \rho \leq \tanh\left(2.0082 + \frac{1.96}{\sqrt{22}}\right)$$

$$0.9202 \leq \rho \leq 0.9845$$

olarak elde edilir.