Exponential Regression Model

Parametric Regression Model:

When a heterogeneous population arises in any practical situation and it is important to consider the relationship of a lifetime to other factors, one way to do this is through a parametric regression model in which the dependence of lifetime on concomitant variables is explicitly recognized.

Exponential Regression Model:

When the individuals have a constant hazard function that may depend on concomitant variables, an exponential regression model is appropriate. The $p.d.f$ of T given x is-

$$f(t \mid x) = \frac{1}{\theta_x} e^{-\frac{t}{\theta_x}} \qquad ; \qquad t > 0 \qquad \cdots \qquad \cdots \qquad \cdots \qquad (i)$$

where x is a vector of regression variables and $\theta_x = E(T \mid x)$.

Various functional forms for θ_x are possible but the most useful is-

$$\theta_x = \exp(X\beta) \qquad \cdots \qquad \cdots \qquad \cdots \qquad (ii)$$

where $X = (x_1,\ x_2,\ \cdots,\ x_p)$ and $\beta = (\beta_1,\ \beta_2,\ \cdots,\ \beta_p)'$ is a vector of regression parameters.

The model (i) is a proportional hazards model. In addition, it can be viewed as a location-scale model for

$$Y = \ln t$$
$$\therefore \quad t = e^Y$$

Now, from the equation (i) and (ii), we have the $p.d.f.$ of Y given X is-

$$f(Y \mid X) = \frac{1}{e^{X\beta}} e^{-\left(e^{Y - X\beta}\right)} e^Y \qquad \text{Hint: } f(y) = f(t^*)|J|, J = \frac{dt}{dy}$$

$$= e^{Y - X\beta} e^{-\left(e^{Y - X\beta}\right)}$$

$$\therefore \quad f(Y \mid X) = \exp\left\{(Y - X\beta) - \exp(Y - X\beta)\right\} \qquad ; \qquad -\infty < Y < \infty$$

Alternatively, let,

$$Y - X\beta = Z$$
$$\Rightarrow \quad Y = X\beta + Z \qquad \cdots \qquad \cdots \qquad \cdots \qquad (iii)$$

where, Z has a standard extreme value distribution with $p.d.f.$

$$f(Z) = \exp[Z - \exp(Z)] \qquad ; \qquad -\infty < Z < \infty$$

Then the regression model expressed in (iii) is called the exponential regression model.

Estimation Procedure of the Exponential Regression Parameters:

Suppose that each individual is associated with each individual's lifetime or censoring time t_i and a regression vector $X_i = (X_{i1},\ X_{i2},\ \cdots,\ X_{ip})$. The notation $i \in D$ and $i \in C$ will be used to refer to individuals i for which t_i is a lifetime and a censoring time respectively.

Since we work with $y_i = \ln t_i$, the $p.d.f.$ of Y given X is-

$$f(Y \mid X) = \exp\{(y - X\beta) - \exp(y - X\beta)\} \qquad ; \qquad -\infty < Y < \infty$$

and the survival time is given by

$$S(Y \mid X) = \exp\{-\exp(y - X\beta)\}$$

The likelihood function for a censored sample based on n individuals is-

$$L(\beta) = \prod_{i \in D} \exp\{(y_i - X_i\beta) - \exp(y_i - X_i\beta)\} \prod_{i \in C} \exp\{-\exp(y_i - X_i\beta)\}$$

Thus, we have

$$\begin{aligned}
\ln L(\beta) &= \sum_{i \in D} \{(y_i - X_i\beta) - \exp(y_i - X_i\beta)\} + \sum_{i \in C} \{-\exp(y_i - X_i\beta)\} \\
&= \sum_{i \in D} (y_i - X_i\beta) - \sum_{i \in D} \exp(y_i - X_i\beta) - \sum_{i \in C} \exp(y_i - X_i\beta)
\end{aligned}$$

$$\therefore \quad \ln L(\beta) = \sum_{i \in D} (y_i - X_i\beta) - \sum_{i=1}^{n} \exp(y_i - X_i\beta)$$

The first and second derivatives of $\ln L(\beta)$ are

$$\frac{\partial \ln L(\beta)}{\partial \beta_r} = -\sum_{i \in D} X_{ir} + \sum_{i=1}^{n} X_{ir} \exp(y_i - X_i\beta) = U_r(\beta) \qquad (say)$$

and

$$\frac{\partial^2 \ln L(\beta)}{\partial \beta_r \partial \beta_s} = -\sum_{i=1}^{n} X_{ir} X_{is} \exp(y_i - X_i\beta) = U_{rs}(\beta) \qquad (say) \quad ; \ r, s = 1(1p)$$

The maximum likelihood equations $\dfrac{\partial \ln L(\beta)}{\partial \beta_r} = 0$ are readily solved by the Newton-Raphson method to get the maximum likelihood estimator of β.

By using the Newton-Raphson method we consider the following steps:

- Obtain an initial estimate β_0 *of* $\hat{\beta}$.
- Calculate $U(\beta_0)$ and $G(\beta_0)$
- Calculate the next approximation β_1 *to* $\hat{\beta}$ using $\beta_1 = \beta_0 - G(\beta_0)^{-1} U(\beta_0)$
- Calculate $U(\beta_1)$, $G(\beta_1)$ and find $\beta_2 = \beta_1 - G(\beta_1)^{-1} U(\beta_1)$. Similarly, $\beta_i = \beta_{i-1} - G(\beta_{i-1})^{-1} U(\beta_{i-1})$

Continue this until convergence is achieved. One can stop when β_0 *and* β_1 are close together and $U(\beta_1)$ is close to 0.

If $|\beta_2 - \beta_1| < \varepsilon$ where ε small quantity then β_2 is MLE of β. Otherwise, we consider the next iteration. In general, if $|\beta_i - \beta_{i-1}| < \varepsilon$ then β_i is MLE of β.

Now, the $p \times p$ observed information matrix is $I_0 = -\left.\frac{\partial^2 \ln L(\beta)}{\partial\beta_r \partial\beta_s}\right]_{\beta=\hat{\beta}}$.

The situations in which a fixed censoring time L_i is known for each individual, the expected information matrix can be calculated as

$$I_{rs} = E\left[-\frac{\partial^2 \ln L(\beta)}{\partial\beta_r \partial\beta_s}\right] \quad = \sum_{i=1}^{n} X_{ir} X_{is}\left[1 - \exp\left(-L_i e^{-X_i\beta}\right)\right];\ r, s = 1,\ 2,\ \cdots,\ p$$

The expected information matrix in the uncensored case is-

$$I_{rs} = \sum_{i=1}^{n} X_{ir} X_{is} \qquad ; \qquad r, s = 1,\ 2,\ \cdots,\ p$$

$$\therefore\ \hat{\beta} \sim N\left[\beta,\ I_{obs}^{-1}(\hat{\beta})\right]$$

When $p = 2$ then $I_{obs}^{-1}(\hat{\beta}) = \begin{bmatrix} V(\beta_1) & \operatorname{cov}(\beta_1,\ \beta_2) \\ \operatorname{cov}(\beta_2,\ \beta_1) & V(\beta_2) \end{bmatrix}$.

Inference Procedure of Regression Parameters:

Let us consider the hypothesis $H_0 : \beta_1 = \beta_1^0$, where β' is partitioned as $(\beta_1,\ \beta_2)'$ and where β_1 is $k \times 1$ $(k < p)$ and β_1^0 is a specified vector.

To test the null hypothesis against $H_1 : \beta_1 \neq \beta_1^0$ we can use the following likelihood ratio statistic:

$$\Lambda = -2\ln\left(\frac{L(\beta_1^0,\ \tilde{\beta}_2)}{L(\hat{\beta}_1,\ \hat{\beta}_2)}\right) \sim \chi_k^2$$

where $\tilde{\beta}_2$ is the MLE of β_2 under H_0, and $\hat{\beta} = (\hat{\beta}_1,\ \hat{\beta}_2)$ is the unconstrained MLE. $\tilde{\beta}_2$ is found by solving $\frac{\partial \ln L(\beta)}{\partial \beta_r} = 0$

Decision Rule: If $\Lambda > \chi^2_{k,\ \alpha\%}$ then we can reject H_0 at $\alpha\%$ level of significance, otherwise we fail to reject H_0.

An alternative statistic for testing H_0 based on the asymptotic normal approximation $\hat{\beta} \sim N\left[\beta,\ I_{obs}^{-1}\right]$ is $\Lambda_1 = \left(\hat{\beta}_1 - \beta_1^0\right)' C_{11}^{-1}\left(\hat{\beta}_1 - \beta_1^0\right) \sim \chi_k^2$ where C_{11} is $k \times k$ and $C = I_0^{-1}$ is partitioned as-

$$C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}.$$

Decision Rule: If $\Lambda_1 > \chi^2_{k,\,\alpha\%}$ then we reject H_0 at $\alpha\%$ level of significance, otherwise we failed to reject H_0.

Again if $\hat{\beta} \sim N\left(\beta,\ I(\hat{\beta})^{-1}\right)$.

In practice, under $H_0 : \beta = 0$, the test statistic is $Z = \dfrac{U(0)}{\sqrt{I(0)}} \sim N(0,\,1)$

$$\text{Where,}\quad U(0) = r_2 - \sum_{i=1}^{k}\left[\frac{d_i n_{2i}}{n_{1i} + n_{2i}}\right], \quad \text{and} \quad I(0) = \sum_{i=1}^{k}\left[\frac{d_i n_{1i} n_{2i}}{(n_{1i} + n_{2i})^2}\right].$$

Decision Rule: If Z lies between ± 1.96, then we may fail to reject H_0, otherwise, we may reject the null hypothesis.

Note: The χ^2 approximation for Λ is typically somewhat better than that for Λ_1 in small and moderate sample size.

Distinguish Between the Exponential Regression Model and the Proportional Hazard Regression Model.

The distinction between the exponential regression model and the proportional hazard regression model is given below:

Exponential Regression Model	**Proportional Hazard Regression Model**
$Y = \beta X + Z \cdots \cdots \cdots (1)$; $-\infty < Y < \infty$, where Z has a standard extreme value distribution with $p.d.f.$ $\exp(Z - \exp(Z))$, the regression model expressed in (1) is called the exponential regression model.	The regression model that uses the proportional hazard function as the dependent variable is called the proportional hazard regression model.
This is a parametric model.	This is a semi-parametric model.
Individuals have constant hazard functions that may depend on concomitant variables.	Different individuals have different hazard functions which are proportional to one another.
Exponential regression models are employed in the analysis of survival data on patients suffering from chronic diseases.	Proportional hazard regression models are used in the Biometrical average. They are also used in an engineering context.

Newton-Raphson Method:

When the derivative of $f(x)$ is a single expression and easily found, the real roots of $f(x)=0$ can be computed rapidly by a process called, the Newton-Raphson method.

Let us denote an approximate value of the desired root and let h denote the correction which must be applied to give the exact value of the root, so that $f(x)=0$ for $x=a+h$ we have $f(a+h)=0$

By Taylor expansion we get,

$$f(a+h)=f(a)+hf'(a)+\frac{h^2}{2}f''(a+\theta h) \quad ; \quad 0\le\theta\le 1$$

$$\Rightarrow \quad 0=f(a)+hf'(a) \qquad [\textit{By ignoring the higher order of } h]$$

$$\Rightarrow \quad h=-\frac{f(a)}{f'(a)}$$

which is the approximate root of a.

Improved values are,

$$a_1=a+h_1=a-\frac{f(a)}{f'(a)}$$

The succeeding approximations are

$$a_2=a_1+h_2=a_1-\frac{f(a_1)}{f'(a_1)}$$

$$-----------------$$

$$a_n=a_{n-1}+h_n=a_{n-1}-\frac{f(a_{n-1})}{f'(a_{n-1})}$$