Sample Questions: Log-Normal Regression

STA312 Spring 2019. Copyright information is at the end of the last page.

1. Let the continuous random variable T have median m. Let $Y = g(T)$, where $g(x)$ is an increasing function. Show that the median of Y is $g(m)$. This is why the median of a log-normal is e^{μ}.

2. Show that the expected value of a log-normal is $e^{\mu+\frac{1}{2}\sigma^2}$. Hint: the moment-generating function of a normal random variable is $e^{\mu t+\frac{1}{2}\sigma^2 t^2}$.

3. Write the log-normal regression model in multiplicative form.

4. For a log-normal regression model, show that if $x_{i,k}$ is increased by c units, $E(t_i)$ is multiplied by $e^{c\beta_k}$.

5. If $x_{i,k}$ is increased by one unit, the median of t_i is multiplied by _______.

6. If $x_{i,k}$ is increased by one unit, the *value* of t_i is multiplied by _______.

7. Write the hazard function of a log-normal regression model in terms of $\Phi(x)$, the cumulative distribution function of a standard normal. Is this a proportional hazards model?

8. Show that in general, if $\widehat{\boldsymbol{\theta}}_n \mathrel{\dot\sim} N_k(\boldsymbol{\theta}, \mathbf{V}_n)$ and $\mathbf{a}$ is a non-zero $k \times 1$ vector of constants, then $W = \mathbf{a}^\top \widehat{\boldsymbol{\theta}}_n \mathrel{\dot\sim} N\left(\mathbf{a}^\top \boldsymbol{\theta},\, \mathbf{a}^\top \mathbf{V}_n\, \mathbf{a}\right)$.

9. What is the parameter vector $\boldsymbol{\theta}$ for a log-normal regression model with $p-1$ explanatory variables?

10. For a log-normal regression model, let $\mathbf{x}_{n+1}$ be a $p \times 1$ vector of explanatory variable values, maybe starting with a 1 for the intercept. A new observation (log failure time) could be written $y_{n+1} = \mathbf{x}^\top \boldsymbol{\beta} + \epsilon_{n+1}$, where $\epsilon_{n+1} \sim N(0, \sigma^2)$, and ϵ_{n+1} is independent of $\epsilon 1, \ldots, \epsilon_n$. It is natural to predict the value of y_{n+1} with the estimated expected value, so $\widehat{y}_{n+1} = \mathbf{x}^\top \widehat{\boldsymbol{\beta}}$.

    Let $\mathbf{V}_n$ denote the $(p+1) \times (p+1)$ asymptotic covariance matrix of the parameter vector. What is the asymptotic distribution of $\widehat{y}_{n+1}$?

11. What is the asymptotic distribution of the error in prediction $y_{n+1} - \widehat{y}_{n+1}$? Justify your answer; include calculation of the expected value and variance.

12. What is the standard error of $y_{n+1} - \widehat{y}_{n+1}$. Remember, a standard error is an *estimated* standard deviation, something that can be computed from sample data.

13. Dividing $y_{n+1} - \widehat{y}_{n+1}$ by its standard error, obtain a Z statistic. What is the asymptotic distribution of Z?

14. Use the Z statistic to obtain a 95% prediction interval for y_{n+1}.
