Online Exam Statistical Modeling/ Generalized Linear Models

(WISM-08, EBB883B05)

Wednesday January 27, 2021, 15:00 - 18:00

RULES AND REMARKS:

- Explain in all cases the reasoning leading to your answer and provide each page with your name and student number.
- Clearly indicate your type of curriculum Mathematics or Econometrics.
- The number of points per question are indicated in a box. The exam consists of 4 questions spread over 4 pages.
- Scan your paper and pen type of work into a set of single files or prepare a single document by type writing which you transform into a single pdf. Next upload your file(s) via the Assignment option into the Nestor system using the special Exam Course.
- Indicate explicitly via the Pledge that the work is of your own, that you received no help from other persons and that you completed the exam as were it a regular paper-and-pen held at the university.
- We wish you success with the completion!

1. **Gamma distribution.** The density of the Gamma distribution is

$$f(y;\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-y\beta},$$

where $y > 0$ and $\Gamma(\alpha)$ is the gamma function. Suppose that the shape parameter α is known and the scale parameter β is the one of interest.

(a) $\boxed{7}$ Show that the Gamma density belongs to the exponential family and use properties of the exponential family to find $E[Y]$ and $\text{Var}[Y]$.
Hint: Use that

$$E[a(Y)] = -\frac{c'(\theta)}{b'(\theta)} \quad \text{and} \quad \text{Var}[a(Y)] = \frac{b''(\theta)c'(\theta) - b'(\theta)c''(\theta)}{(b'(\theta))^3}.$$

(b) $\boxed{3}$ Derive the expected value and variance of the Chi-squared distribution with density

$$f(y;p) = \frac{1}{\Gamma(p/2)2^{p/2}} y^{(p/2)-1} e^{-y/2}.$$

2. **The family of exponential distributions** is characterized by the class of density functions

$$f(y;\theta) = \exp\{a(y)b(\theta) + c(\theta) + d(y)\},$$

where the sufficient conditions to interchange the order of differentiation and integration hold. Suppose the observable variables $Y_1, \cdots, Y_n$ are independently distributed with density $f(y_i;\theta_i)$, for $i = 1, \cdots, n$, from a member of the exponential family with $a(y) = y$, $E[Y_i] = \mu_i$ and the link function $g(\mu_i) = \tilde{x}_i^T\beta$, where $\tilde{x}_i^T$ is the row vector with observed values of the explanatory variables for unit i, and β the weights of the generalized linear model.

(a) 10 Show that element jk of the information matrix $\boldsymbol{J}$, can be written as $\boldsymbol{x}_j^T \boldsymbol{W} \boldsymbol{x}_k$ for a certain diagonal matrix $\boldsymbol{W}$, where $\boldsymbol{x}_j$ is column j of the matrix $\boldsymbol{X}$.

(b) 10 The m-th updating step of the Newton-Rapson algorithm can be written as

$$\boldsymbol{b}_{m+1} = \boldsymbol{b}_m + \boldsymbol{J}_m^{-1}\boldsymbol{U}_m,$$

where $\boldsymbol{J}_m$ is the information matrix and $\boldsymbol{U}_m$ the score vector at step m. Recall that the latter is defined as the first order derivative of the log-likelihood function with respect to the parameters. Show that m-th updating step can be written as the solution to a weighted least squares problem leading to the m-th updating step written as

$$\boldsymbol{b}^{(m+1)} = (\boldsymbol{X}^T\boldsymbol{W}^{(m)}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{W}^{(m)}\boldsymbol{z}^{(m)},$$

(c) 5 Give expressions for updating the weights $\boldsymbol{W}$ as well as the term $\boldsymbol{z}$ for Poisson regression using the log link.
Hint 1: Y_i Poisson distributed implies that $P(Y_i = y) = \frac{\mu_i^y e^{-\mu_i}}{y!}$, $E[Y_i] = \text{Var}[Y_i] = \mu_i$.
Hint 2: Recall that

$$w_{ii} = \frac{1}{\text{Var}Y_i}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2, \text{ and } z_i = \boldsymbol{x}_i^T\boldsymbol{b} + (y_i - \mu_i)\left(\frac{\partial\eta_i}{\partial\mu_i}\right)$$

4. **Survival Analysis**. Suppose that Y is a non-negative continuous random variable indicating the survival time with distribution function $F(y)$ and survivor function $S(y) = P(Y \geq y)$.

(a) [10] The hazard function $h(y)$ is defined as the probability of death in $[y, y + \delta y]$ given survival up to y, i.e. $Y > y$, relative to an infinitely small interval ($\delta y \to 0$). Show that

$$h(y) = -\frac{d}{dy} \log[S(y)].$$

(b) [5] Show that the cumulative hazard function, defined as

$$H(y) = \int_0^y h(t)dt,$$

equals the expression $-\log(1 - F(y))$, where F is the distribution function of Y.

(c) [5] The log-logistic distribution with probability density function

$$f(y; \theta, \lambda) = \frac{e^\theta \lambda y^{\lambda-1}}{(1 + e^\theta y^\lambda)^2}$$

is sometimes used for modeling survival times. Find its survival function $S(y; \theta, \lambda)$, hazard function, cumulative hazard function and its median survival time.

3. **Linear Regression.** Consider the linear model $\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{e}$ with fixed $N \times p$ design matrix $\boldsymbol{X}$ and $\boldsymbol{e} \sim N(\boldsymbol{o}, \sigma^2 I)$, with density

$$f(\mathbf{e}) = (2\pi)^{-N/2} |\det \sigma^2 I|^{-1/2} \exp\{-\mathbf{e}^T (\sigma^2 I)^{-1} \mathbf{e}/2\}.$$

Let $\boldsymbol{y}$ be the realized values of $\boldsymbol{Y}$. Let $\|\cdot\|$ be the Euclidian norm.

(a) [5] Show that the maximum likelihood estimator for $\boldsymbol{\beta}$ is $\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}$.

(b) [5] Show that

$$\widehat{\sigma}^2_{ML} = \frac{1}{n}\|\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\|^2,$$

is the maximum likelihood estimator for σ^2 and derive its expected value.

(c) [5] Let $\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T$ be the projection or hat matrix with h_{ik} as its ik-th element. Furthermore, let $\widehat{y}_i = \boldsymbol{x}_i^T\widehat{\boldsymbol{\beta}}$, where $\boldsymbol{x}_i^T$ is the i-th row of $\boldsymbol{X}$. Show that the following properties hold:

$$\mathrm{Var}(\widehat{y}_i) = h_{ii}\sigma^2,$$

$$\mathrm{Cov}(\widehat{y}_i, \widehat{y}_k) = h_{ik}\sigma^2,$$

$$\mathrm{Var}(y_i - \widehat{y}_i) = (1 - h_{ii})\sigma^2,$$

$$\mathrm{Cov}(\widehat{y}_i, y_k - \widehat{y}_k) = 0, \text{ for all } i, k,$$

(d) [10] Show that $h_{ii} \in [0, 1]$ for all i, and that $h = \max_i h_{ii} \to 0$, as $n \to \infty$, is sufficient for $\widehat{y}_i$ to be consistent in tending in probability to $\boldsymbol{x}_i^T\boldsymbol{\beta}$ and also for $\widehat{\boldsymbol{\beta}}$ to tend in probability to $\boldsymbol{\beta}$.
Hint: It is sufficient to show that the variance tends to zero.

(e) [10] Show that the residuals $\widehat{\boldsymbol{e}} = \boldsymbol{y} - \widehat{\boldsymbol{y}}$ and the predicted values $\widehat{\boldsymbol{y}}$, defined as $\widehat{\boldsymbol{y}} = \boldsymbol{X}\widehat{\boldsymbol{\beta}}$, are stochastically independent and use this to show that $\widehat{\boldsymbol{\beta}}$ and $\widehat{\sigma}^2_{ML}$ are stochastically independent.