STAT 613 Final (2 hours) **May 7th, 2012**

There are 3 questions, you need to answer all of them.

Good Luck and have a great summer

(1) Let us suppose that we observe the response variable and regressor (Y_t, X_t). Y_t and X_t are related through the model

$$Y_t = g(X_t) + \varepsilon_t$$

where ε_t are iid Gaussian random variables (with mean zero and variance σ^2) which are independent of the regressors X_t. X_t are independent random variables, and the density of X_t is f. Suppose that it is wrongly assumed that Y_t satisfies the model $Y_t = \beta X_t + \varepsilon_t$, where ε_t are iid Gaussian random variables (with mean zero and variance σ^2, which can be assumed known). [15]

- (i) Given $\{(Y_t, X_t)\}_{t=1}^T$, what is the maximum likelihood estimator of β?
- (ii) Derive an expression for the limit of this estimator (ie. what is the misspecified likelihood estimator actually estimating).
- (iii) Derive an expression for the Kullbach-Leibler information between the true model and the best fitting misspecified model (that you derived in part (ii)).

(2) Suppose that Y_t are discrete random variables with mean $\mathbb{E}(Y_t|X_t) = \lambda_t$ and variance $\mathrm{var}(Y_t|X_t) = \lambda_t(1 + \xi\lambda_t)$ (ie. there is dispersion). Suppose we model $\lambda_t = \exp(\beta_0 X_t)$, where X_t are observed univariate regressors.

The solution of the estimating equation

$$\sum_{t=1}^{T}(Y_t - \exp(\beta X_t))X_t = 0,$$

is used as an estimator of β_0. [15]

- (i) Using the delta method it can be shown that $\sqrt{T}(\hat{\beta}_T - \beta_0) \xrightarrow{D} \mathcal{N}(0, V)$. Derive an expression for V_T.
- (ii) Suppose the random variables X_t are normally distributed (standard normal with mean zero and variance one). Find a expression for V_T in terms of β and ξ.

  You may want to use that

  $$\frac{1}{\sqrt{2pi}}\int_{-\infty}^{\infty} \exp(\frac{-1}{2}(x-\mu)^2)dx = 1.$$

- (iii) Having obtained estimators for β (and thus the mean λ_t), suggest a method for estimating ξ.

(3) Let us suppose that T_i are independent, identically distributed survival times, with density $f(t;\theta_0)$ and survivial function $\mathcal{F}(t;\theta_0)$. The parameter θ_0 is unknown (assume it does not lie on the support of f).

We do not always observe T_i and sometimes censoring is believed to occur. In this case, the censoring variable δ_i and the censored time Y_i, (where $T_i \geq Y_i$) are observed. To estimate θ_0, we use the regular censored likelihood

$$\mathcal{L}_T(\theta) = \sum_i \delta_i \log f(T_i;\theta) + \sum_i (1-\delta_i)\log \mathcal{F}(Y_i;\theta),$$

where δ_i is an indicator variable to denote censoring, $\delta_i = 1$ when it is believed that a variable has not be censored and $\delta_i = 0$ when it has been censoring. Let $\hat{\theta}_T$ denote the parameter which maximises $\mathcal{L}_T(\theta)$. [20]

(i) Suppose that the censoring variable δ_i is misleading. δ_i can be zero or one at random, independent of T_i; and $P(\delta_i = 1) = \pi$ and $P(\delta_i = 0) = 1-\pi$. Thus $\mathbb{E}[(1-\delta_i)\log\mathcal{F}(Y_i;\theta)] = \mathbb{E}(1-\delta_i)\mathbb{E}(\log\mathcal{F}(Y_i;\theta))$. Moreover, the Y_i that is observed is actually $Y_i = \min(T_i, C_i)$, where C_i is another survival time (denote it as G) that is independent of T_i and δ_i.

Show that in the case $\{C_i\}$ are iid random variables, it is unlikely that $\hat{\theta}_T$ is estimating θ_0 (it is not a consistent estimator of θ_0).

(ii) Show that in the special case that C_i also has the survival density $f(x;\theta_0)$, then $\hat{\theta}_T$ is a consistent estimator of θ_0.

Amendment: This actually turns out not to be the case! See my solutions.

(iii) In the special case described in (ii) show that $\mathbb{E}(\frac{\partial \mathcal{L}_T(\theta)}{\partial\theta}\rfloor_{\theta_0})^2 = -\mathbb{E}(\frac{\partial^2 \mathcal{L}_T(\theta)}{\partial\theta^2}\rfloor_{\theta_0})$.

Note: This will only be true in the case that $\mathbb{E}(\frac{\partial \mathcal{L}_T(\theta)}{\partial\theta}\rfloor_{\theta_0}) = 0$.

(iv) Since the survival times are randomly being censored (where the censoring is independent of the survival time), another way to estimate θ_0 is to restrict ourselves to the non-censored values. That is use the parameter which maximises

$$\mathcal{L}_{2,T}(\theta) = \sum_i \delta_i \log f(T_i;\theta),$$

as an estimator of θ. By considering $\mathbb{E}\frac{\partial \mathcal{L}_T(\theta)}{\partial\theta})^2$ and $\mathbb{E}(\frac{\partial \mathcal{L}_{2T}(\theta)}{\partial\theta})^2$ explain why in the special case that the density of C_i is $f(x;\theta_0)$ we 'gain' by using the censored likelihood $\mathcal{L}_T(\theta)$ to estimate θ_0.