2016 Qual, Section 2

Ann Marie Weideman

1

1. (25 points) Suppose that $y_1, \ldots, y_n$ are positive and independent random variables, where

$$p(y_i|\mu_i) = \frac{1}{\mu_i}\exp(-y_i/\mu_i), \quad \mu_i > 0, \tag{1}$$

where $E(y_i|\mu_i) = \mu_i$, $i = 1, \ldots n$. Let $\theta_i = 1/\mu_i$.

(a) (3 points) Suppose that θ_i is random with $\theta_i \sim \text{Gamma}(a_i, b_i)$, where $a_i/b_i = \exp(-x_i'\beta)$ and $a_i = 3$. Further assume $\text{Var}(\theta_i) = \tau\exp(x_i'\beta)$. Here, x_i is a $p \times 1$ vector of covariates and β is a $p \times 1$ vector of regression coefficients, and β is unknown. Derive the **marginal** mean and variance of y_i, that is, compute $E(y_i)$ and $\text{Var}(y_i)$.

(b) (3 points) Under the same assumptions as part (a), derive the marginal distribution of y_i.

(c) (7 points) Under the same assumptions as part (a), derive the score test for testing $H_0 : \tau = 0$ and give its asymptotic distribution under the null hypothesis.

(d) Now suppose we take μ_i to be a **fixed and unknown parameter** and we incorporate over-dispersion by taking $\text{Var}(y_i) = \sigma^2(v_i + \mu_i)$ where v_i is the variance function of the GLM in (1). Let $\mu_i = \exp\{x_i'\beta\}$.

(i) (5 points) Derive the quasi-likelihood score equations for β and a moment estimator for σ^2.

(ii) (7 points) Let $\hat{\beta}_P$ denotes the quasi-likelihood estimate of β. Derive the asymptotic covariance matrix for $\hat{\beta}_P$.

AMW

2016 Qual Section 2

1. Suppose $Y_1, \ldots, Y_n$ are positive and independent RVs where

$$p(y_i | \mu_i) = \frac{1}{\mu_i} \exp(-y_i/\mu_i), \quad \mu_i > 0$$

where $E(y_i | \mu_i) = \mu_i$, $i = 1, \ldots, n$. Let $\theta_i = \frac{1}{\mu_i}$

a) Suppose that θ_i is random w/ $\theta_i \sim \text{Gamma}(a_i, b_i)$ where $a_i/b_i = \exp(-x_i'\beta)$ and $a_i = 3$.

Further, assume $Var(\theta_i) = \frac{1}{3} \cdot \exp(x_i'\beta)$. Here X_i is a $p \times 1$ vector of covariates and β is a $p \times 1$ vector of regression coefficients and β is unknown.

$\boxed{\text{Compute } E[y_i]}$

$$E[y_i] = E[\underbrace{E(y_i | \mu_i)}_{\mu_i}] = E[\mu_i] = E\left[\frac{1}{\theta_i}\right]$$

Given $\theta_i \sim \text{Gamma}(a_i, b_i)$,

$$\text{Then, } E\left[\frac{1}{\theta_i}\right] = \int_0^\infty \frac{1}{\theta_i} \cdot \frac{b_i^{a_i}}{\Gamma(a_i)} \theta_i^{a_i - 1} e^{-b_i\theta_i} \, d\theta_i = \int_0^\infty \frac{b_i^{a_i}}{\Gamma(a_i)} \theta_i^{(a_i-1)-1} e^{-b_i\theta_i} \, d\theta_i$$

$$= \frac{b_i^{a_i}}{\Gamma(a_i)} \cdot \frac{\Gamma(a_i - 1)}{b_i^{a_i - 1}} = \frac{b_i}{a_i - 1} = \boxed{\frac{1}{2} b_i}$$

$\boxed{\text{Compute } Var[y_i]}$

$$Var[y_i] = E[\underbrace{Var[y_i | \mu_i]}_{\substack{\mu_i^2 \\ \text{b/c } y_i|\mu_i \sim Exp(\mu_i)}}] + Var[\underbrace{E[y_i | \mu_i]}_{\mu_i}] = E[\mu_i^2] + Var[\mu_i] = E[\mu_i^2] + (E[\mu_i^2] - E[\mu_i]^2)$$

$$= 2E[\mu_i^2] - \underbrace{E[\mu_i]}_{(\frac{1}{2} b_i \text{ from previous})}^2 = 2E[\mu_i^2] - \frac{1}{4} b_i^2$$

where

$$E[\mu_i^2] = E\left[\frac{1}{\theta_i^2}\right] = \int_0^\infty \frac{1}{\theta_i^2} \cdot \frac{b_i^{a_i}}{\Gamma(a_i)} \cdot \theta_i^{a_i - 1} e^{-b_i\theta_i} \, d\theta_i = \int_0^\infty \frac{b_i^{a_i}}{\Gamma(a_i)} \cdot \theta_i^{(a_i-2)-1} e^{-b_i\theta_i} \, d\theta_i$$

$$= \frac{b_i^{a_i}}{\Gamma(a_i)} \cdot \frac{\Gamma(a_i - 2)}{b_i^{a_i - 2}} = \frac{b_i^2}{(\underset{3}{\underset{\shortparallel}{a_i}} - 1)(\underset{3}{\underset{\shortparallel}{a_i}} - 2)} = \frac{b_i^2}{2}$$

$$\Rightarrow Var[y_i] = 2\left(\frac{b_i^2}{2}\right) - \frac{1}{4} b_i^2 = b_i^2 - \frac{b_i^2}{4} = \frac{4b_i^2 - 1b_i^2}{4} = \boxed{\frac{3b_i^2}{4}}$$

2016 Qual Section 2 AMW

1.b) Under the same assumptions as a), derive the marginal distribution of y_i

$\ulcorner$ Know $p(y_i|\mu_i) = \frac{1}{\mu_i}\exp(-y_i/\mu_i)$, $\mu_i > 0$

$\Rightarrow p(y_i|\theta_i) = \theta_i \exp(-\theta_i y_i)$, $\theta_i > 0$

and $\theta_i \sim Gamma(a_i, b_i)$

$\Rightarrow p(y_i|\theta_i) = \frac{p(y_i,\theta_i)}{p(\theta_i)} \Rightarrow p(y_i,\theta_i) = p(y_i|\theta_i)p(\theta_i)$ by Bayes rule.

Then, $p(y_i) = \int_{\theta_i} p(y_i,\theta_i)\,d\theta_i = \int_0^\infty \theta_i e^{-\theta_i y_i} \cdot \frac{b_i^{a_i}}{\Gamma(a_i)} \theta_i^{a_i-1} e^{-b_i\theta_i}\,d\theta_i$

$$= \int_0^\infty \frac{b_i^{a_i}}{\Gamma(a_i)} \theta_i^{(a_i+1)-1} e^{-(b_i+y_i)\theta_i}\,d\theta_i = \frac{b_i^{a_i}}{\cancel{\Gamma(a_i)}} \cdot \frac{\overbrace{\Gamma(a_i+1)}^{a_i\cancel{\Gamma(a_i)}}}{(b_i+y_i)^{a_i+1}} = \frac{a_i b_i^{a_i}}{(b_i+y_i)^{a_i+1}}$$

$$= \boxed{\frac{3b_i^3}{(b_i+y_i)^4}}, \quad y_i > 0. \lrcorner$$

AMW

2016 Qual, Section 2

1 c) Under the same assumptions as in a), derive the score test for testing $H_0: \tau = 0$ and give its asymptotic distn. under H_0.

Want to test $H_0: \tau = 0$ vs. $H_0: \tau > 0$

Have $E(\theta_i) = K(x_i'\beta)$, $Var(\theta_i) = \tau \exp(x_i'\beta)$

General form of variance is $Var(\theta_i) = \tau f_i(x_i'\beta) \Rightarrow f_i = \exp(x_i'\beta)$

① Write in exponential family form.

$$p(y_i|\mu_i) = \frac{1}{\mu_i}\exp\left\{-\frac{1}{\mu_i}y_i\right\} = \exp\left\{-\frac{1}{\mu_i}y_i + \log\left(\frac{1}{\mu_i}\right)\right\} = \underbrace{\exp\left\{-\frac{1}{\mu_i}y_i - \log(\mu_i)\right\}}$$

Form: $\exp\{\phi(\theta y - b(\theta) - c(y)) - \tfrac{1}{2}S(y,\phi)\}$

where $\phi = 1$

$\theta_i = \frac{1}{\mu_i}$

$y = -y_i$ (I e-mailed Dr. Ibrahim, and he said to put the negative on the y_i in order to get the canonical parameter to match the parameterization given earlier in the problem.)

$$b(\theta_i) = \log(\mu_i) = \log\left(\frac{1}{\theta_i}\right) = -\log(\theta_i) \Rightarrow \dot{b}(\theta) = -\frac{1}{\theta_i}$$

$$\Rightarrow \text{Since } \underbrace{\theta_i = x_i'\beta}_{\text{by defn.}} \Rightarrow \underbrace{\mu_i}_{\text{mean function}} = \dot{b}(\theta_i) = -\frac{1}{(x_i'\beta)} = -(x_i'\beta)^{-1}$$

Thus, $\mu_i = -(x_i'\beta)^{-1}$ (mean function based on canonical link for an exponential distribution - checked on wiki)

And, $\dot{b}(\theta) = -\frac{1}{\theta_i} \Rightarrow \ddot{b}(\theta_i) = \frac{1}{\theta_i^2} \Rightarrow \dddot{b}(\theta_i) = \frac{-2}{\theta_i^3} \Rightarrow b^{(4)}(\theta_i) = \frac{6}{\theta_i^4}$

② Write the general form for score test of $H_0: \tau = 0$

From slide 903, $S_\tau = \left.\frac{\partial_\tau \ln(\alpha)^2}{\sigma_\tau^2} \cdot 1\{\partial_\tau \ln(\alpha) > 0\}\right|_{\tilde{\alpha}}$ where $\tilde{\alpha}$ is the estimate of α under the null $H_0: \tau = 0$

$$\xrightarrow{d} \underbrace{0.5\chi_0^2 + 0.5\chi_1^2}_{\text{(a mixture of chi-squares)}}$$

Memorize by recalling that

$$SC_n = \left.\underbrace{\dot{\ell}(\alpha)'}_{\partial_\tau \ln(\alpha)}\underbrace{[I(\alpha)]^{-1}}_{\sigma_\tau^2}\underbrace{\dot{\ell}(\alpha)}_{\partial_\tau \ln(\alpha)}\right|_{\tilde{\alpha}}$$

cont'd →

AMW

1 c) cont'd

③ Numerator of S_τ :

$$\partial_\tau \ln(\alpha) = \frac{1}{2}\sum_{i=1}^{n} f_i \{ (y_i-\mu_i)^2 - \ddot{b}(\cdot)\} = \frac{1}{2}\sum_i e^{x_i'\beta} \{ (y_i-\mu_i)^2 - \mu_i^2\} = \frac{1}{2}\sum_i e^{x_i'\beta}\{ y_i^2 - 2\mu_i y_i\}$$

$$= \frac{1}{2}\sum_i e^{x_i'\beta} \{ y_i^2 + 2(x_i'\beta)^{-1} y_i\}$$

$$\Rightarrow \partial_\tau \ln(\alpha) = \frac{1}{4}\left[\sum_i e^{x_i'\beta}\{ y_i^2 + 2(x_i'\beta)^{-1} y_i\}\right]^2$$

④ Denominator of S_τ : Have $\sigma_\tau^2 = I_{\tau\tau} - I_{\tau\beta}' I_{\beta\beta}^{-1} I_{\tau\beta}$

(Memorize by recalling that MVN variance is $Var[X_{want} \mid X_{given}]$
$= \Sigma_{want} - \underbrace{\Sigma_{b/w}'}_{\text{between}} \Sigma_{given}^{-1} \Sigma_{b/w}$.)

where $I_{\tau\tau} = \frac{1}{4}\sum_i f_i^2 \{ 2\ddot{b}(\theta_i)^2 + b^{(4)}(\theta_i)\} = \frac{1}{2\cancel{4}}\sum_i e^{x_i'\beta}\{\cancel{2}(x_i'\beta)^{-4} + \cancel{6}^{3}(x_i'\beta)^{-4}\}$

$$= \frac{1}{2}\sum_i e^{x_i'\beta} \{ (x_i'\beta)^{-4} + 3(x_i'\beta)^{-4}\} = 2\sum_i e^{x_i'\beta} / (x_i'\beta)^4$$

$I_{\beta\tau} = D_\theta(\beta)' W_2 \cdot 1_n$ where $D_\theta(\beta) = \frac{\partial\theta}{\partial\beta} = \underbrace{diag(x_i'\beta) X}_{\text{derivative of } \theta_i = x_i'\beta \text{ w.r.t. } \beta}$.

and $W_2 = diag\left(\frac{1}{2} f_i \dddot{b}(\theta_i)\right) = diag\left(\frac{1}{\cancel{2}} e^{x_i'\beta} \cdot \frac{-\cancel{2}}{(x_i'\beta)^3}\right)$

$$= diag\left(- e^{x_i'\beta} / (x_i'\beta)^3\right)$$

and $I_{\beta\beta} = D_\theta(\beta)' W_1 \cdot D_\theta(\beta)$ where $D_\theta(\beta) = diag(x_i'\beta) X$ (as found above)

and $W_1 = diag(\ddot{b}(\theta_i)) = diag((x_i'\beta)^{-2})$

④ Conclude Thus, reject H_0 if $S_\tau > 0.5\, \chi_0^2(1-\alpha) + 0.5\, \chi_1^2(1-\alpha)$. ┘

1.d) Now suppose we take μ_i to be a fixed and unknown parameter and we incorporate over-dispersion by taking $Var(y_i) = \sigma^2(v_i + \mu_i)$ where v_i is the variance function of the GLM in (1). Let $\mu_i = \exp\{x_i'\beta\}$,

(i) Derive the quasi-likelihood score eqns. for β and a moment estimator for σ^2.

Q: What is quasi-likelihood? What is the point? Who gives a shit?

A: • Quasi-likelihood estimation is one way of allowing for overdispersion.

• So, if we want to perform a score test to make inference about β, it is important that we account for this overdispersion.

• Quasi-likelihood provides an important method for making statistical inference w/out making parametric assumptions.

• Quasi-likelihood can be applied to indep & dep observations.

The quasi log-likelihood for μ based on the data y is given by

$$\boxed{l_q(\mu, y) = \sum_{i=1}^{n} q_i(\mu_i, y_i)}, \quad \underbrace{q_i(\mu_i, y_i)}_{\text{behaves like a log-likelihood}} = \int_{y_i}^{\mu_i} \frac{(y_i - t)}{\sigma^2 V_i(t)}\, dt$$

MEMORIZE

$$\Rightarrow l_q(\mu, y) = \sum_i \int_{y_i}^{\mu_i} \frac{(y_i - t)}{\sigma^2 V_i(t)}\, dt \Rightarrow \frac{\partial l_q}{\partial \beta} = \sum_i \frac{\partial \mu_i}{\partial \beta}\left(\frac{y_i - \mu_i}{\sigma^2 V_i(\mu_i)} - \frac{y_i - y_i}{\sigma^2 V_i(y_i)}\right)$$ By FTOC

$$= \sum_i \frac{\partial \mu_i}{\partial \beta}\left(\frac{y_i - \mu_i}{\sigma^2 V_i(\mu_i)}\right) = \sum_i \exp\{x_i'\beta\} x_i \left(\frac{y_i - \mu_i}{\sigma^2 V_i(\mu_i)}\right) = (\sigma^2)^{-1} D^T V^{-1} e \Rightarrow D^T V^{-1} e = 0$$

where $D = diag(\exp\{x_i'\beta\} x_i)$ for $i = 1, \dots, n$, $V = diag(V_i) = diag(\tau \exp\{x_i'\beta\})$ and $e = y - \mu$.

ii) Derive the moment estimator of σ^2.

Since $E\left[\sum_{i=1}^{n} \frac{(y_i - \mu_i)^2}{V_i(\mu_i)}\right] = n\sigma^2$, then subbing $\hat{\mu}_i = \exp\{x_i'\hat{\beta}\}$ and $\hat{V}_i = \tau \exp\{x_i'\hat{\beta}\} = \tau \hat{\mu}_i$

Then, $\hat{\sigma}^2 = \frac{1}{n-p} \sum_{i=1}^{n} \frac{(y_i - \hat{\mu}_i)^2}{\tau \hat{\mu}_i}$ for $\hat{\mu}_i = \exp\{x_i'\hat{\beta}\}$.

Correct for df lost due to p parameters

AMW

1 d) ii) Let $\hat{\beta}_P$ denote the quasi-likelihood of β. Derive the asymptotic covariance matrix for $\hat{\beta}_P$.

Let $S_n(\beta)$ be the score eqn.

Let β_* be the true value of β.

approximate since we drop off ~~terms~~ derivatives of order 2 and higher

Taylor expansion about $\underbrace{\hat{\beta}_P = \beta^*}_{x=a}$ gives: $0 = \overbrace{S_n(\hat{\beta}_P)}^{f(x)} \approx \underbrace{S_n(\beta_*)}_{f(a)} + \underbrace{\partial_\beta S_n(\beta_*)}_{f'(a)} \underbrace{(\hat{\beta}_P - \beta_*)}_{x-a}$

just gorgeous?

$\Rightarrow -S_n(\beta_*) = \partial_\beta S_n(\beta_*)(\hat{\beta}_P - \beta_*) \Rightarrow (\hat{\beta}_P - \beta_*) = [-\partial_\beta S_n(\beta_*)]^{-1} S_n(\beta_*)$

$\Rightarrow \sqrt{n}(\hat{\beta}_P - \beta_*) = \left[-\frac{1}{n}\partial_\beta S_n(\beta_*)\right]^{-1} \cdot \frac{1}{\sqrt{n}} S_n(\beta_*)$

Know $\mathrm{Cov}(\sqrt{n}\,\hat{\beta}_P) \approx \underbrace{\left[-\frac{1}{n}\partial_\beta S_n(\beta_*)\right]^{-1}}_{\text{find}} \underbrace{\left(\frac{1}{\sqrt{n}} S_n(\beta_*)\right)}_{\text{find}} \left[-\frac{1}{n}\partial_\beta S_n(\beta_*)\right]^{-1}$

expected value of $e_i(\beta_*) = 0$

<u>Where</u> $-\frac{1}{n}\partial_\beta S_n(\beta_*) = -\frac{1}{n}\partial_\beta\left[\sum_i \frac{\partial\mu_i}{\partial\beta} V_i(\beta_*)^{-1} \overbrace{e_i(\beta_*)}^{= y_i - \mu_i(\beta)}\right]$

product rule $= -\frac{1}{n}\left\{\sum_i \partial_\beta\left[\frac{\partial\mu_i}{\partial\beta} V_i(\beta_*)^{-1}\right] e_i(\beta_*) - \sum_i \frac{\partial\mu_i}{\partial\beta} V_i(\beta_*)^{-1} \frac{\partial\mu_i'}{\partial\beta}\right\}$

Then, by WLLN (i.e., $\bar{X}_n \xrightarrow{p} \mu$), $-\frac{1}{n}\partial_\beta S_n(\beta_*) \xrightarrow{p} -\frac{1}{n} E\left\{\sum_i \partial_\beta\left[\frac{\partial\mu_i}{\partial\beta} V_i(\beta_*)^{-1}\right] e_i(\beta_*) - \sum_i \frac{\partial\mu_i}{\partial\beta} V_i(\beta_*)^{-1} \frac{\partial\mu_i'}{\partial\beta}\right\}$

$= \frac{1}{n}\sum_i \frac{\partial\mu_i}{\partial\beta} V_i(\beta_*)^{-1} \frac{\partial\mu_i'}{\partial\beta}$

Thus, $-\frac{1}{n}\partial_\beta S_n(\beta_*) \approx \frac{1}{n} D^T V^{-1} D$

and <u>where</u> $\mathrm{Cov}\left[\frac{1}{\sqrt{n}} S_n(\beta_*)\right] = \mathrm{Cov}\left[\frac{1}{\sqrt{n}}\sum_i \frac{\partial\mu_i}{\partial\beta} V_i^{-1}(\beta_*) \overbrace{e_i(\beta_*)}^{\mathrm{Cov}(e_i(\beta_*)) = \sigma^2 V_i(\beta_*)}\right]$

$= \frac{1}{n}\sum_i \mathrm{Cov}\left(\frac{\partial\mu_i}{\partial\beta} V_i^{-1}(\beta_*) e_i(\beta_*)\right) = \frac{1}{n}\sum_i \frac{\partial\mu_i}{\partial\beta}\mathrm{Cov}\left(V_i^{-1}(\beta_*) e_i(\beta_*)\right)\frac{d\mu_i'}{d\beta}$

$= \frac{1}{n}\sum_i \frac{\partial\mu_i}{\partial\beta} V_i^{-2}(\beta_*)\, \sigma^2 V_i(\beta_*) \frac{\partial\mu_i'}{\partial\beta} = \frac{\sigma^2}{n}\sum_i \frac{\partial\mu_i}{\partial\beta} V_i(\beta_*) \frac{\partial\mu_i'}{\partial\beta}$

$= \sigma^2 D^T V^{-1} D / n$

$\Rightarrow \mathrm{Cov}(\sqrt{n}\,\hat{\beta}_P) \approx \left(\frac{1}{n} D^T V^{-1} D\right)^{-1} \left(\sigma^2 D^T V^{-1} D / n\right) \left(\frac{1}{n} D^T V^{-1} D\right)^{-1} = \sigma^2 n (D^T V^{-1} D)^{-1}$

$\Rightarrow \mathrm{cov}(\hat{\beta}_P) \approx \hat{\sigma}^2 \left(D_\mu(\hat{\beta}_P)' V(\hat{\mu})^{-1} D_\mu(\hat{\beta}_P)\right)^{-1}$

where $\hat{\sigma}^2$ is as derived in part i), $\underbrace{\hat{\beta}_P = \mathrm{argmax}\ I_q(\mu(\beta), y)}_{\text{Slide 413}}$, and $\hat{\mu}_i = \exp\{x_i'\hat{\beta}_P\}$