# Extra Practice Questions

## 1 Questions

**Q1:** Suppose $\sqrt{n}\left(\hat{\beta}_n - \beta\right) \xrightarrow{d} \mathcal{N}(0, V)$ for some positive definite matrix $V$. Let $f : \mathbb{R}^{k+1} \to \mathbb{R}$ be continuously differentiable and nonzero at $\beta$ with derivative $f'(\beta)$. Let $\hat{V}_n$ be a consistent estimator of $V$.

a) Derive the limiting distribution of $\sqrt{n}\left(f\left(\hat{\beta}_n\right) - f(\beta)\right)$.

b) Construct a test of asymptotic size $\alpha$ of

$$H_0 : f(\beta) \leq 0 \quad vs. \quad H_1 : f(\beta) > 0.$$

Justify your answer. Show that the power of the test converges to 1 if $f(\beta) > 0$ and to 0 if $f(\beta) < 0$.

c) Use your answer to a) and b) to construct a confidence region $C_n$ of asymptotic level $1 - \alpha$ for $f(\beta)$, and show that as $n \to \infty$,

$$\mathrm{P}_\beta\left(f(\beta) \in C_n\right) \to 1 - \alpha.$$

di) Let $\beta = (\beta_0, \beta_1, \beta_2) \in \mathbb{R}^3$. Explain how to conduct a test of the null hypothesis that $\beta_0 + \beta_1 - \beta_2 \leq 3$ vs. the alternative that $\beta_0 + \beta_1 - \beta_2 > 3$ of asymptotic level $\alpha$.

dii) Explain how to conduct a test of the null hypothesis that $\beta_0 + \beta_1 = \beta_1 + \beta_2 = 3$ vs. the alternative $H_1$ : Not $H_0$ of asymptotic size $\alpha$.

**Q2:** Consider the linear model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u, \quad \mathrm{E}(u|x) = 0, Var(u|x) = \sigma^2.$$

You estimate $\beta = (\beta_0, \beta_1, \beta_2)$ by OLS. You wish to test $H_0 : c'\beta = 0$ vs. $H_1 : c'\beta \neq 0$, where $c = (c_0, c_1, c_2)$ and $c_2 \neq 0$, but you do not want to compute a matrix product. In this question we will show how to obtain the t-statistic for $c'\beta$ in standard regression packages by rewriting the regression equation. The idea is to redefine the regressors appropriately to produce the t-statistic for a linear combination of parameters directly. Try it! For $j = 0, 1$, let $\delta_j = \beta_j$ and let $\delta_2 = c'\beta$.

i) Rewrite $y = x'\beta + u$ as $y = \tilde{x}'\delta + u$. What is $\tilde{x}$?

ii) Show that $\delta = G\beta$, where

$$G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ c_0 & c_1 & c_2 \end{pmatrix}.$$

iii) Show that

$$G^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{c_0}{c_2} & -\frac{c_1}{c_2} & \frac{1}{c_2} \end{pmatrix}$$

by checking that $G^{-1}G = GG^{-1} = I_3$.

iv) Stack the variables into matrices $Y, X, \tilde{X}, U$. Show that $\tilde{X} = XG^{-1}$.

v) Show that an OLS regression of $Y$ on $\tilde{X}$ produces $\hat{\delta} = G\hat{\beta}$. (This takes care of the numerator of the t-stat).

vi) Show that the SSR of a regression of $Y$ on $X$ is the same as a regression of $Y$ on $\tilde{X}$. Conclude that the estimate of $\sigma^2$ is identical in both regressions.

vii) Show that the estimated variance covariance matrix of $\hat{\delta}$ is given by

$$\hat{Var}\left(\hat{\delta}|X\right) = G\left[\hat{\sigma}^2 \left(X'X\right)^{-1}\right]G'.$$

viii) Conclude that the "t-value" for the coefficient $\delta_2$ computed by your regression package is equivalent to the test statistic you would use to test $H_0$ from Week 8.

**Q3:** Let $X_i$, $i = 1, 2, \ldots, n$, be an iid sequence of random variables with common pdf

$$f_\theta(x) = \begin{cases} \frac{\theta c^\theta}{x^{\theta+1}} & \text{if } c < x \\ 0 & \text{otherwise} \end{cases}$$

Here, $c > 0$ is known and $\theta > 0$ is the unknown parameter of interest.

a) Write down the likelihood function.

b) Derive the maximum likelihood estimator $\hat{\theta}_n$ for $\theta$.

c) Show that $\hat{\theta}_n$ is a consistent estimator of $\theta$. Is $\hat{\theta}_n$ an unbiased estimator of $\theta$? Explain.

d) Derive the (nondegenerate) limiting distribution of $\tau_n\left(\hat{\theta}_n - \theta\right)$ for an appropriate choice of $\tau_n$.

**Q4:** Consider the following causal model for wages:

$$\log(y_T) = \gamma + x'\beta + u,$$
$$\log(y_{NT}) = x'\beta + u,$$

where

$$y_T = \text{wage if employed in tech industry,}$$

$$y_{NT} = \text{wage if employed outside tech industry,}$$
$$x = \text{observed determinants of wage.}$$

The observed wage is $y = y_T d + (1-d)\, y_{NT}$, where

$$d = \text{indicator for tech industry job.}$$

Let $(y_i, d_i, x_i)_{i=1}^n$ be an iid sample drawn from the distribution of $(y, d, x)$. Define $w = (d, x')'$. Assume that $x$ includes a constant, that $\mathbb{E}[wu] = 0$, that both $\mathbb{E}[ww']$ and $\operatorname{Var}[wu]$ exist, and that $\mathbb{E}[ww']$ is invertible. Define $\delta = \exp(\gamma) - 1$.

a) Interpret $\delta$.

b) Suggest a consistent estimator $\hat{\delta}_n$ of $\delta$. Justify your answer.

c) Show that $\sqrt{n}\left(\hat{\delta}_n - \delta\right) \xrightarrow{\mathrm{d}} N(0, \tau^2)$ for some $\tau^2$ as $n \to \infty$.

d) Suggest a consistent estimator $\hat{\tau}_n^2$ of $\tau^2$. Justify your answer.

e) Construct an asymptotic $1-\alpha$ two-sided confidence region for $\delta$. Justify your answer.

## 2 Solutions

**Q1:** a) By the delta method:

$$\sqrt{n}\left(f\left(\hat{\beta}_n\right) - f(\beta)\right) \xrightarrow{d} \mathcal{N}\left(0, f'(\beta)\, V\, [f'(\beta)]^T\right).$$

b) First note that $\sqrt{n}\left(\hat{\beta}_n - \beta\right) \xrightarrow{d} \mathcal{N}(0, V)$ implies $\hat{\beta}_n \xrightarrow{p} \beta$. Since $f'$ is continuous, $f'\left(\hat{\beta}_n\right) \xrightarrow{p} f'(\beta)$ by the continuous mapping theorem. Since

$$\begin{pmatrix} \left[f'\left(\hat{\beta}_n\right)\right]^T \\ vec\left(\hat{V}_n\right) \end{pmatrix} \xrightarrow{p} \begin{pmatrix} [f'(\beta)]^T \\ vec(V) \end{pmatrix},$$

it follows again by the continuous mapping theorem that $f'\left(\hat{\beta}_n\right) \hat{V}_n \left[f'\left(\hat{\beta}_n\right)\right]^T \xrightarrow{p} f'(\beta)\, V\, [f'(\beta)]^T$. By Slutsky's theorem, we obtain

$$\frac{\sqrt{n}\left(f\left(\hat{\beta}_n\right) - f(\beta)\right)}{\sqrt{f'\left(\hat{\beta}_n\right) \hat{V}_n\, [f'(\beta)]^T}} \xrightarrow{d} \mathcal{N}(0, 1). \tag{1}$$

We propose the test statistic

$$T_n = \frac{\sqrt{n} f\left(\hat{\beta}_n\right)}{\sqrt{f'\left(\hat{\beta}_n\right) \hat{V}_n\, [f'(\beta)]^T}},$$

and the test

$$\phi_n = \mathbf{1}\left(T_n > z_{1-\alpha}\right).$$

If $f(\beta) < 0$, then

$$\frac{T_n}{\sqrt{n}} - \frac{z_{1-\alpha}}{\sqrt{n}} \xrightarrow{p} \frac{f(\beta)}{\sqrt{f'(\beta) V [f'(\beta)]^T}} - 0 < 0,$$

so the limit is a negative constant. Since 0 is a continuity point of this limiting distribution,

$$\mathrm{P}\left(T_n > z_{1-\alpha}\right) = 1 - \mathrm{P}\left(\frac{T_n}{\sqrt{n}} - \frac{z_{1-\alpha}}{\sqrt{n}} \leq 0\right) \to 1 - 1 = 0.$$

A similar argument provides that if $f(\beta) > 0$, $\mathrm{P}\left(T_n > z_{1-\alpha}\right) \to 1$. If $f(\beta) = 0$, then result (3) provides that $T_n \xrightarrow{d} \mathcal{N}(0,1)$. Since $z_{1-\alpha}$ is a continuity point of $\mathcal{N}(0,1)$,

$$\begin{aligned}
\mathrm{P}\left(T_n > z_{1-\alpha}\right) &= 1 - \mathrm{P}\left(T_n \leq z_{1-\alpha}\right) \\
&= 1 - (1-\alpha) = \alpha.
\end{aligned}$$

c) Result (3) implies that

$$\mathrm{P}\left(\frac{\sqrt{n}\left(f\left(\hat{\beta}_n\right) - f(\beta)\right)}{\sqrt{f'\left(\hat{\beta}_n\right)\hat{V}_n [f'(\beta)]^T}} \leq z_{1-\alpha}\right) \to 1 - \alpha,$$

because the limiting distribution of the asymptotic pivot is $\mathcal{N}(0,1)$. Rearranging yields that

$$C_n = \left[f\left(\hat{\beta}_n\right) - z_{1-\alpha}\frac{\sqrt{f'\left(\hat{\beta}_n\right)\hat{V}_n [f'(\beta)]^T}}{\sqrt{n}}, \infty\right)$$

is a $1-\alpha$ confidence set for $f(\beta)$.

di) Let $r = [1, 1, -1]$. Using previous results gives

$$\frac{\sqrt{n}\left(r'\hat{\beta}_n - r'\beta\right)}{\sqrt{r'\hat{V}_n r}} \xrightarrow{d} \mathcal{N}(0,1).$$

Rewriting $H_0$ as $\beta_0 + \beta_1 - \beta_2 - 3 \leq 0$ suggests the test statistic

$$T_n = \frac{\sqrt{n}\left(r'\hat{\beta}_n - 3\right)}{\sqrt{r'\hat{V}_n r}}.$$

The corresponding test and its properties were derived in part b).

dii) This is a test of two linear restrictions. Let

$$R = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}; \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}; c = \begin{pmatrix} 3 \\ 3 \end{pmatrix}.$$

We may rewrite the null and alternative hypotheses as $H_0 : R\beta = c$ vs. $H_1 : R\beta \neq c$. A result from Week 9 provides that under $H_0$:

$$n \cdot \left(R\hat{\beta}_n - c\right)' \left(R\hat{V}_n R'\right)^{-1} \left(R\hat{\beta}_n - c\right) \xrightarrow{d} \chi_2^2,$$

and so we reject iff $T_n > \chi^2_{2,1-\alpha}$, where the latter quantity is the $1-\alpha$ quantile of the $\chi_2^2$ distribution. This test has asymptotic size $\alpha$ by construction.

**Q2:** i) Rearranging variables gives

$$y = \beta_0 \left[1 - \frac{c_0}{c_2} x_2\right] + \beta_1 \left[x_1 - \frac{c_1}{c_2} x_2\right] + (c_0\beta_0 + c_1\beta_1 + c_2\beta_2) \frac{x_2}{c_2} + u,$$

so

$$\tilde{x} = \left(\left[1 - \frac{c_0}{c_2} x_2\right], \left[x_1 - \frac{c_1}{c_2} x_2\right], \frac{x_2}{c_2}\right)$$

ii) We find

$$G\beta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ c_0 & c_1 & c_2 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ c_0\beta_0 + c_1\beta_1 + c_2\beta_2 \end{pmatrix} = \delta.$$

iv) We compute

$$XG^{-1} = \begin{pmatrix} | & | & | \\ X_0 & X_1 & X_2 \\ | & | & | \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{c_0}{c_2} & -\frac{c_1}{c_2} & \frac{1}{c_2} \end{pmatrix} = \begin{pmatrix} | & | & | \\ X_0 - \frac{c_0}{c_2} X_2 & X_1 - \frac{c_0}{c_2} X_2 & \frac{1}{c_2} X_2 \\ | & | & | \end{pmatrix} = \tilde{X}.$$

v) From part iv), a regression of $Y$ on $\tilde{X}$ is equivalent to running OLS on

$$Y = XG^{-1}\delta + U,$$

which produces

$$\begin{aligned} \hat{\delta} &= \left(\left[G^{-1}\right]' X'XG^{-1}\right)^{-1} \left[G^{-1}\right]' X'Y \\ &= G\left(X'X\right)^{-1} G' \left[G^{-1}\right]' X'Y \\ &= G\left(X'X\right)^{-1} X'Y \\ &= G\hat{\beta}. \end{aligned}$$

vi) Since $\hat{\beta} = G^{-1}\hat{\delta}$, we have

$$\left(Y - XG^{-1}\hat{\delta}\right)'\left(Y - XG^{-1}\hat{\delta}\right) = \left(Y - X\hat{\beta}\right)'\left(Y - X\hat{\beta}\right) = SSR.$$

Since both regressions have the same number of regressors they produce the same estimate of $\sigma^2$.

vii) The estimated variance matrix is

$$\hat{Var}\left(\hat{\delta}|X\right) = \hat{\sigma}^2\left(\left[G^{-1}\right]' X'XG^{-1}\right)^{-1} = G\left[\hat{\sigma}^2\left(X'X\right)^{-1}\right]G'.$$

viii) The standard error of $\hat{\delta}_2$ computed from the regression package is the $(3,3)$ entry of $G\left[\hat{\sigma}^2\left(X'X\right)^{-1}\right]G'$, which is $c'\left[\hat{\sigma}^2\left(X'X\right)^{-1}\right]c$. The t-stat is therefore

$$\frac{\hat{\delta}_2}{se\left(\hat{\delta}_2\right)} = \frac{c_0\hat{\beta}_0 + c_1\hat{\beta}_1 + c_2\hat{\beta}_2}{\sqrt{c'\left[\hat{\sigma}^2\left(X'X\right)^{-1}\right]c}} = \frac{\sqrt{n}\left(c'\hat{\beta} - 0\right)}{\sqrt{c'\left[\hat{\sigma}^2\left(\frac{X'X}{n}\right)^{-1}\right]c}},$$

which is the statistic we would use if assuming homoskedasticity.

**Q3:** a) Since observations are iid, the likelihood function is given by

$$\ell_n(\theta) = \prod_{i=1}^{n} f_\theta(X_i) = \prod_{i=1}^{n}\left(\frac{\theta c^\theta}{x_i^{\theta+1}}\right)\mathbf{1}(x_i > c)$$

b) Taking logs and normalizing, for all $x_i > c$, we obtain:

$$L_n(\theta) = \frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{\theta c^\theta}{X_i^{\theta+1}}\right) = \log(\theta) + \theta\log c - (\theta+1)\frac{1}{n}\sum_{i=1}^{n}\log X_i.$$

Maximizing $L_n(\theta)$ with respect to $\theta$ gives the following first-order condition:

$$\frac{1}{\hat{\theta}_n} + \log c - \frac{1}{n}\sum_{i=1}^{n}\log X_i = 0 \quad \Rightarrow \hat{\theta}_n = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\log X_i - \log c}.$$

c) Since the $X_i$ are iid, if the mean exists, then the WLLN gives us that: $\frac{1}{n}\sum_{i=1}^{n}\log X_i \xrightarrow{p} \mathbb{E}\left[\log X_i\right]$. We then use integration by parts to derive $\mathbb{E}\left[\log X_i\right]$.

$$\begin{aligned}
\mathbb{E}\left[\log X_i\right] &= \int_c^\infty \log x f_\theta(x)\,dx = \int_c^\infty \log x\frac{\theta c^\theta}{x^{\theta+1}}dx \\
&= \left[\left(-c^\theta x^{-\theta}\right)\log x\right]_c^\infty - \int_c^\infty\left(\frac{1}{x}\right)\left(-c^\theta x^{-\theta}\right)dx \\
&= \log c + \int_c^\infty c^\theta x^{-\theta-1}dx
\end{aligned}$$

$$
\begin{aligned}
&= \log c + \left[-\frac{c^\theta}{\theta} x^{-\theta}\right]_c^\infty \\
&= \log c + \frac{1}{\theta} < \infty,
\end{aligned}
$$

where we used L'Hopital's rule to obtain:

$$
\lim_{x\to\infty} \frac{\log x}{x^\theta} = \lim_{x\to\infty} \frac{\frac{1}{x}}{\theta x^{\theta-1}} = \lim_{x\to\infty} \frac{1}{\theta x^\theta} = 0.
$$

Let us define $g(x) \coloneqq \frac{1}{x-\log c}$, which is continuous everywhere except at $x = \log c$. Noting that $\theta > 0$ so that $\mathbb{E}[\log X_i] \neq \log c$, we apply the CMT to obtain:

$$
\hat{\theta}_n = g\left(\frac{1}{n}\sum_{i=1}^n \log X_i\right) \xrightarrow{\text{p}} g(\mathbb{E}[\log X_i]) = \theta.
$$

Therefore, $\hat{\theta}_n$ is consistent for $\theta$. However, $\hat{\theta}_n$ is biased. To see why, note that $\hat{\theta}_n$ is a strictly convex function of $\frac{1}{n}\sum_{i=1}^n \log X_i$. So, by Jensen's Inequality, we have:

$$
\theta = \frac{1}{\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n \log X_i\right] - \log c} < \mathbb{E}\left[\hat{\theta}_n\right] = \mathbb{E}\left[\frac{1}{\frac{1}{n}\sum_{i=1}^n \log X_i - \log c}\right],
$$

since $\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n \log X_i\right] = \mathbb{E}[\log X_i] = \log c + \frac{1}{\theta}$.

d) By the Central Limit Theorem, if the variance exists, we know that:

$$
\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n \log X_i - \mathbb{E}[\log X_i]\right) \xrightarrow{\text{d}} N(0, \text{Var}[\log X_i]),
$$

Since $\text{Var}[\log(X_i)] = \mathbb{E}\left[(\log X_i)^2\right] - \mathbb{E}[\log(X_i)]^2$, we use integration by parts to find $\mathbb{E}\left[(\log X_i)^2\right]$.

$$
\begin{aligned}
\mathbb{E}\left[(\log X_i)^2\right] &= \int_c^\infty (\log(x))^2 \frac{\theta c^\theta}{x^{\theta+1}} dx \\
&= \left[\left(-c^\theta x^{-\theta}\right)(\log(x))^2\right]_c^\infty - \int_c^\infty \left(2\log(x)\frac{1}{x}\right)\left(-c^\theta x^{-\theta}\right) dx \\
&= (\log c)^2 + 2c^\theta \int_c^\infty \log(x)\, x^{-\theta-1} dx \\
&= (\log c)^2 + 2\frac{1}{\theta}\int_c^\infty \log(x) \frac{\theta c^\theta}{x^{\theta+1}} dx \\
&= (\log c)^2 + 2\frac{1}{\theta}\mathbb{E}[\log X_i] \\
&= (\log c)^2 + \frac{2}{\theta}\log c + \frac{2}{\theta^2}.
\end{aligned}
$$

Then, using our expression for $\mathbb{E}[\log X_i]$, we write:

$$\operatorname{Var}[\log(X_i)] = (\log c)^2 + \frac{2}{\theta}\log c + \frac{2}{\theta^2} - \left(\log c + \frac{1}{\theta}\right)^2 = \frac{1}{\theta^2}.$$

Hence:

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\log X_i - \frac{1}{\theta}\right) \xrightarrow{\mathrm{d}} N\left(0, \frac{1}{\theta^2}\right).$$

Let $g(x) := \frac{1}{x-\log c}$ and $g'(x) = -\frac{1}{(x-\log c)^2}$. By the Delta Method, we obtain:

$$\sqrt{n}\left(g\left(\frac{1}{n}\sum_{i=1}^{n}\log X_i\right) - g\left(\frac{1}{\theta} + \log c\right)\right) \quad \Rightarrow \sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow{\mathrm{d}} -\theta^2 N\left(0, \frac{1}{\theta^2}\right) \stackrel{\mathrm{d}}{=} N\left(0, \theta^2\right).$$

**Q4:** a) Since $y = y_{NT} + d\,(y_T - y_{NT})$, where $y_T$ denotes the individual's potential wage with a tech job and $y_{NT}$ denotes the potential wage outside of tech, we have

$$\gamma = \log(y_T) - \log(y_{NT}) \implies \delta = \frac{y_T - y_{NT}}{y_{NT}},$$

so $100 \times \delta$ represents the percentage by which wages in the tech industry exceed wages not in tech holding other covariates and $u$ fixed. It is not simply "the percentage difference in wage between tech jobs and non-tech jobs" because if instead $d = 1$ for jobs outside tech, $\gamma$ would switch sign. By the same construction, we would obtain

$$\delta = \exp(\gamma) - 1 = \frac{y_{NT} - y_T}{y_T}.$$

The denominator is now $y_T$, reflecting the fact that if the base group changes, the percentage difference relative to the base group depends on whether the base group has higher or lower wages to begin with. The new interpretation would be the percentage by which the wage outside tech would exceed the wage in tech holding other covariates fixed. Note that when $\gamma \approx 0$, $\exp(\gamma) - 1 \approx \gamma$, which leads to the same conclusion no matter what the base group is (except for the sign change), and yields the common interpretation of a log-level model like this one that $100 \times \gamma$ represents the ceteris paribus percentage difference in wage resulting from a switch to a job in tech.

b) First construct a consistent estimator $\hat{\gamma}_n$ for $\gamma$. Since $d$ and $x$ are uncorrelated with the error term, the OLS estimator is suitable:

$$\begin{pmatrix} \hat{\gamma}_n \\ \hat{\beta}_n \end{pmatrix} = \left(\frac{1}{n}\sum_{i=1}^{n} w_i w_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} w_i \log y_i\right) \xrightarrow{p} \begin{pmatrix} \gamma \\ \beta \end{pmatrix}.$$

c) We have

$$\begin{pmatrix} \hat{\gamma}_n \\ \hat{\beta}_n \end{pmatrix} = \left(\frac{1}{n}\sum_{i=1}^{n} w_i w_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} w_i \log y_i\right)$$

$$\Rightarrow \sqrt{n}\begin{pmatrix} \hat{\gamma}_n - \gamma \\ \hat{\beta}_n - \beta \end{pmatrix} = \left(\frac{1}{n}\sum_{i=1}^{n} w_i w_i'\right)^{-1}\left(\sqrt{n}\frac{1}{n}\sum_{i=1}^{n} w_i u_i\right)$$

By WLLN and the Continuous Mapping Theorem, we know:

$$\left(\frac{1}{n}\sum_{i=1}^{n} w_i w_i'\right)^{-1} \xrightarrow{\text{p}} \mathbb{E}\left[ww'\right]^{-1}.$$

By Central Limit Theorem and Slutsky's Theorem, we have that:

$$\left(\sqrt{n}\frac{1}{n}\sum_{i=1}^{n} w_i u_i\right) \xrightarrow{\text{d}} N\left(0, \text{Var}\left[wu\right]\right) \quad \Rightarrow \sqrt{n}\begin{pmatrix} \hat{\gamma}_n - \gamma \\ \hat{\beta}_n - \beta \end{pmatrix} \xrightarrow{\text{d}} N\left(0, \Omega\right),$$

where $\Omega = \mathbb{E}\left[ww'\right]^{-1} \text{Var}\left[wu\right] \mathbb{E}\left[ww'\right]^{-1}$. Thus, by the Delta method, setting $g\left(b\right) = \exp\left(b\right) - 1$, we have:

$$\sqrt{n}\left(\left(\exp\left(\hat{\gamma}_n\right) - 1\right) - \left(\exp\left(\gamma\right) - 1\right)\right) \xrightarrow{\text{d}} N\left(0, \exp\left(2\gamma\right)\Omega_{11}\right).$$

We conclude that $\sqrt{n}\left(\hat{\delta}_n - \delta\right) \xrightarrow{\text{d}} N\left(0, \tau^2\right)$, where $\tau^2 = \exp\left(2\gamma\right)\Omega_{11}$.

d) A consistent estimator is $\hat{\tau}_n^2 = \exp\left(2\hat{\gamma}_n\right)\hat{\Omega}_{n,11}$,where:

$$\hat{\Omega}_n = \left(\frac{1}{n}\sum_{i=1}^{n} w_i w_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} w_i w_i' \hat{u}_i^2\right)\left(\frac{1}{n}\sum_{i=1}^{n} w_i w_i'\right)^{-1}$$

By the CMT, $\exp\left(2\hat{\gamma}_n\right)$ is consistent. Applying the CMT again yields that $\hat{\tau}_n^2$ is consistent.

e) Note that $\frac{\sqrt{n}\left(\hat{\delta}_n - \delta\right)}{\sqrt{\hat{\tau}_n^2}} \xrightarrow{\text{d}} N\left(0, 1\right)$. Therefore, the confidence region is given by

$$C_n = \left\{c \in \mathbb{R} : \frac{\sqrt{n}\left|\hat{\delta}_n - c\right|}{\sqrt{\hat{\tau}_n^2}} \leq z_{1-\frac{\alpha}{2}}\right\} = \left[\hat{\delta}_n - z_{1-\frac{\alpha}{2}}\frac{\sqrt{\hat{\tau}_n^2}}{\sqrt{n}}, \hat{\delta}_n + z_{1-\frac{\alpha}{2}}\frac{\sqrt{\hat{\tau}_n^2}}{\sqrt{n}}\right],$$

where $z_{1-\frac{\alpha}{2}}$ is the $1-\alpha/2$ quantile of standard normal distribution. Note that:

$$\begin{aligned}
\mathbb{P}\left(\delta \in C_n\right) &= \mathbb{P}\left(\hat{\delta}_n - z_{1-\frac{\alpha}{2}}\frac{\sqrt{\hat{\tau}_n^2}}{\sqrt{n}} \leq \delta \leq \hat{\delta}_n + z_{1-\frac{\alpha}{2}}\frac{\sqrt{\hat{\tau}_n^2}}{\sqrt{n}}\right) \\
&= \mathbb{P}\left(\left|\hat{\delta}_n - \delta\right| \leq z_{1-\frac{\alpha}{2}}\frac{\sqrt{\hat{\tau}_n^2}}{\sqrt{n}}\right) \\
&= \mathbb{P}\left(\frac{\sqrt{n}\left|\hat{\delta}_n - \delta\right|}{\sqrt{\hat{\tau}_n^2}} \leq z_{1-\frac{\alpha}{2}}\right) \\
&\to \Phi\left(z_{1-\frac{\alpha}{2}}\right) - \Phi\left(-z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha.
\end{aligned}$$