Math 582 Quiz 8 2022 Name__________

pm 1) Let $\boldsymbol{w}_B \sim N_p\left(\mathbf{0}, \frac{\boldsymbol{\Sigma}}{B}\right)$. Then $\boldsymbol{w}_B \xrightarrow{D} \boldsymbol{w}$ as $B \to \infty$. Find $\boldsymbol{w}$.

$\underset{\sim}{W} \sim N_p(\underset{\sim}{0}, 0)$ $P(\underset{\sim}{W} = 0) = 1$ point mass at $\underset{\sim}{0}$

20

2) Let $\boldsymbol{w}_n^* \sim N_p(\mathbf{0}, \boldsymbol{\Sigma}_n)$. If $\boldsymbol{\Sigma}_n \xrightarrow{P} \boldsymbol{\Sigma}$, then $\boldsymbol{w}_n^* \xrightarrow{D} \boldsymbol{w}$ as $n \to \infty$. Find $\boldsymbol{w}$.

(Σ_n and Σ are constant wrt bootstrap dist)

$\underset{\sim}{W} \sim N_p(\underset{\sim}{0}, \Sigma)$

20

3) Let $x_1, ..., x_n$ be iid with mean $E(x) = \mu$ and variance $V(x) = \sigma^2 > 0$. ~~Show that~~ Then $\sum_{i=1}^n (x_i - \overline{x}_n)^2 = \sum_{i=1}^n (x_i - \mu + \mu - \overline{x}_n)^2 = \sum_{i=1}^n (x_i - \mu)^2 - n(\overline{x} - \mu)^2$.

pm

10 a) $\frac{1}{n}\sum_{i=1}^n (x_i - \mu)^2 \xrightarrow{P} \theta$. What is θ?

$\theta = E(x_i - \mu)^2 = \boxed{\sigma^2 = V(x)}$

by WLLN

10 b) $n(\overline{x} - \mu)^2 = [\sqrt{n}(\overline{x} - \mu)]^2 \xrightarrow{D} W$. What is W?

F d z z

$\sqrt{n}(\bar{x} - \mu) \to N(0, \sigma^2) = Z \sim \sigma\, N(0, 1)$ by CLT

so $W = Z^2 \sim \sigma^2 \chi^2_1$

Continuous mapping th

20

4) The table below shows simulation results for bootstrapping OLS (reg) and forward selection (vs) with C_p when $\boldsymbol{\beta} = (1, 1, 0, 0)^T$. The β_i columns give coverage = the proportion of CIs that contained β_i and the average length of the CI. The test is for $H_0 : (\beta_3, \beta_4)^T = \mathbf{0}$ and H_0 is true. The "coverage" is the proportion of times the prediction region method bootstrap test failed to reject H_0. Since 1000 runs were used, a cov in [0.93,0.97] is reasonable for a nominal value of 0.95. Output is given for three different error distributions. If the coverage for both methods ≥ 0.93, the method with the shorter average CI length was more precise. (If one method had coverage ≥ 0.93 and the other had coverage < 0.93, we will say the method with coverage ≥ 0.93 was more precise.)

a) For β_2, β_3, and β_4, which method, forward selection or the OLS full model, was more precise?

forward selection   8.9 < 10.6, 2.7 < 3.2, 4.6 < 5.4

Table 1: Bootstrapping Forward Selection, $n = 100, p = 4, \psi = 0.9, B = 1000$

		β_1	β_2	β_3	β_4	test
reg	cov	0.93	0.95	0.95	0.94	0.95
	len	1.266	10.703	10.666	10.650	2.547
vs	cov	0.95	0.93	0.997	0.995	0.989
	len	1.260	8.901	8.986	8.977	2.759
reg	cov	0.94	0.93	0.95	0.94	0.95
	len	0.393	3.285	3.266	3.279	2.475
vs	cov	0.94	0.97	0.998	0.997	0.995
	len	0.394	2.773	2.721	2.733	2.703
reg	cov	0.95	0.94	0.95	0.95	0.95
	len	0.656	5.493	5.465	5.427	2.493
vs	cov	0.93	0.95	0.998	0.998	0.977
	len	0.657	4.599	4.655	4.642	2.783

b) The test "length" is the average length of the interval $[0, D_{(U_B)}] = D_{(U_B)}$ where the test fails to reject H_0 if $D_{\mathbf{0}} \leq D_{(U_B)}$. The OLS full model is asymptotically normal, and hence for large enough n and B the reg len row for the test column should be near $\sqrt{\chi^2_{2,0.95}} = 2.477$.

Were the three values in the test column for reg within 0.1 of 2.477?

Yes

$2.477 \pm .1 = [2.377, 2.577]$

$2.547, 2.475, 2.493$

40