Problem Honors

1.- Sea $X_1, \dots, X_n$ una muestra aleatoria de una v.a. Bernoulli con $P(X=1)=p$, donde $p \in (0,1)$ es desconocido. Sea $\hat{\theta}$ el estimador de Máxima verosimilitud de $\theta = p(1-p)$.

a) Muestre que $\hat{\theta}$ es asintóticamente normal cuando $p \neq \frac{1}{2}$.
Respuesta.
Teorema 1. (Diapositivas, 72p)
Bajo apropiadas condiciones de regularidad, el mle $\hat{\theta}$

$$\sqrt{n}\,(\hat{\theta} - \theta) \xrightarrow{d} N(0, I_1(\theta)^{-1}).$$

Teorema 2. (Casella, 499p).

$$Var(h(\hat{\theta}) \mid \theta) \approx \frac{[h'(\theta)]^2}{I_n(\theta)} \simeq \frac{[h'(\theta)]^2 \big|_{\theta=\hat{\theta}}}{-\frac{\partial^2}{\partial \theta^2} \log L(\theta|x) \big|_{\theta=\hat{\theta}}}.$$

Por las clases sabemos que el estimador de MLE para el caso Bernoulli de p es $\hat{p} = \bar{X}$. Ahora, esto implica que el estimador de la varianza para el caso Bernoulli es

$$\hat{\theta} = \hat{p}(1-\hat{p}) = \bar{x}(1-\bar{x}).$$

Una manera de abordar este problema sería calcular las varianza de $\bar{X}(1-\bar{X})$ pero eso complica la respuesta. Por lo que utilizaremos el teorema 2 para encontrarla.

$$\widehat{Var}(\hat{\theta}^n) = \widehat{var}\left(\hat{p}^n(1-\hat{p}^n)\right) = \frac{\left[\frac{\partial}{\partial p}\left(p(1-p)\right)\right]^2\Big|_{p=\hat{p}^n}}{-\frac{\partial^2}{\partial p^2}\log L(p|x)\Big|_{p=\hat{p}^n}}$$

$$= \frac{(1-2p)^2\Big|_{p=\hat{p}^n}}{\frac{n}{p(1-p)}\Big|_{p=\hat{p}^n}}$$

$$= \frac{\hat{p}(1-\hat{p})(1-2\hat{p}^n)^2}{n}$$

Observemos que si $p=\frac{1}{2} \Rightarrow Var(\hat{\theta}^n)=0$, por lo que se estará sobreestimando. Por lo que consideraremos el caso cuando $p\neq\frac{1}{2}$. Por lo tanto, ocupando el teorema 1 podemos concluir que

$$\sqrt{n}\left(\hat{\theta}^n - p(1-p)\right) \xrightarrow{d} N\left(0, (1-p)p(1-2p)^2\right) \Leftrightarrow \sqrt{n}\left(\bar{x}(1-\bar{x}) - p(1-p)\right) \xrightarrow{d} N\left(0, p(1-p)(1-2p)^2\right)$$

Es decir, $\hat{\theta}$ es asintoticamente normal (para $p\neq\frac{1}{2}$).

b) Cuando $p=\frac{1}{2}$, usando una normalización adecuada, deriva una distribución asintótica no degenerada de $\hat{\theta}$.

Respuesta.

Ocuparemos el teorema 5.5.26 del Casella.

Teorema S.S. 26

Sea Y_n una m.a. que satisface $\sqrt{n}(Y_n - \theta) \to N(0, \sigma^2)$ en distribución. Para alguna función g y un valor especifico de θ, supongace que $g'(\theta) = 0$ y $g''(\theta)$ existe y no es 0. Entonces

$$n\left[g(Y_n) - g(\theta)\right] \xrightarrow{d} \sigma^2 \frac{g''(\theta)}{2} \chi_1^2 .$$

En el inciso anterior y probamos que se satisface

$$\sqrt{n}(Y_n - \theta) \xrightarrow{d} N(0, \sigma^2). \; \star, \; Y_n = \bar{X}.$$

Sea $g(\theta) = p(1-p) \Rightarrow g'(\theta) = 1 - 2p$

Pero como $p = \frac{1}{2} \Rightarrow$

$$g(\theta) = \frac{1}{4} \quad \& \quad g'(\theta) = 0.$$

Y además $g''(\theta) = -2.$

Por lo tanto podemos concluir que

$$n\left[Y_n(1-Y_n) - \frac{1}{4}\right] \xrightarrow{d} -\frac{1}{4} \chi_1^2$$

Es decir, encontramos una distribución asintotica no degenerada.

Nota: $\star$ no es la demostración es directa considerando que $\hat{\theta}_{MLE} = \bar{X} \Rightarrow \sqrt{n}(Y_n - p) \xrightarrow{d} N(0, p(1-p))$, es decir, usar la media.

3. Para el caso presentado en el ejemplo anterior sin usar el teorema de Wilks, justifique que la distribución asintótica de $-2 \log \lambda(x)$ es χ_1^2.

Respuesta:

Enunciemos los siguientes teoremas:

Teorema 1

Sea $\{X_n\}$ convergente a X en distribución, y sea $\{Y_n\}$ convergente en probabilidad a cero. Entonces $\{X_n Y_n\}$ converge en probabilidad a cero.

Teorema 2.

Sea $\hat{\theta}^{(n)}$ el estimador de MLE de $\overset{\circ}{\theta}$, si este existe y es único (si no, $\hat{\theta}^n$ se define algún arbitrario). Bajo ciertas condiciones, el límite distribucional de $\sqrt{n}(\hat{\theta}^n - \overset{\circ}{\theta})$ es $N(0, J^{-1})$, y $\hat{\theta}^n$ converge en probabilidad a $\overset{\circ}{\theta}$.

Teorema 3.

Sea $\{\theta^{(n)}\}$ convergente en probabilidad a θ. Bajo ciertas condiciones para $i,j = 1, 2, \ldots, r$, de la secuencia

$$\left\{-n^{-1} \frac{\partial^2 L_n(\theta^{(n)})}{\partial\theta_i \partial\theta_j}\right\}$$ converge en probabilidad a J_{ij}.

Teorema 4.

$$-2\lambda(x) = \sqrt{n}\,(\hat{\theta}^n - \overset{\circ}{\theta})^T \left[-n^{-1} \frac{\partial^2 L_n(\theta^*)}{\partial\theta_i \partial\theta_j}\right] \sqrt{n}(\hat{\theta}^n - \overset{\circ}{\theta}).$$

Ocupando el teorema 4 tenemos que

$$-2\log \lambda(x) = \sqrt{n}\,(\hat{\theta}^n - \overset{\circ}{\theta})^T \left[-n^{-1} \frac{\partial^2 l_n(\theta^*)}{\partial\theta_i \partial\theta_j}\right] \sqrt{n}\,(\hat{\theta}^n - \overset{\circ}{\theta}).$$

Y ocupando el teorema 2, tenemos que $\{\theta^*\}$ converge en probabilidad a θ°. Y ocupando el teorema 3, tenemos que para cada $i, j = 1, \ldots, r$ de los elementos de la matriz $\left[-n^{-1} \frac{\partial^2 L_n(\theta^*)}{\partial\theta_i \partial\theta_j}\right]$ converge a J_{ij}. Por lo que, tenemos que

$$-2\log \lambda(x) = \sqrt{n}\,(\theta^n - \overset{\circ}{\theta})^T J \sqrt{n}\,(\theta^n - \overset{\circ}{\theta}).$$

Y por último ocupanto nuevamente el teorema 2, podemos ver que $\sqrt{n}\,(\theta^n - \overset{\circ}{\theta}) \sim N(0, J^{-1})$ y por las propiedades de la distribución Normal

$\Rightarrow$

$$-2\log \lambda(x) = \sqrt{n}\,(\theta^n - \overset{\circ}{\theta})^T J \sqrt{n}\,(\theta^n - \overset{\circ}{\theta}) \sim \chi_1^2$$