CAPÍTULO 1:

- Un estadístico es una función de las observaciones muestrales, ej: $g(X_1, \dots, X_n) = \frac{1}{n}\sum_{i=1}^{n} X_i$.

- Parámetros poblacionales
  - media $= \mu = \frac{1}{N}\sum_{i=1}^{N} X_i$
  - varianza $= \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i-\mu)^2$
  - proporción $= p = \frac{X}{N} = \frac{\text{éxitos}}{\text{pruebas}}$.

- Estadísticos
  - media $\bar{X} = \frac{1}{n}\sum X_i$
  - varianza $S^2 = \frac{1}{n-1}\sum(X_i-\bar{X})^2$.
  - proporción $p_X = \frac{X}{n}$.

lemma $\mathbb{E}[\bar{X}] = \mu$, $\operatorname{var}(\bar{X}) = \sigma^2/n$.

- La desviación típica de $\bar{X} = \sigma/\sqrt{n}$.

- Teorema.
  - $X_i \sim N(\mu,\sigma)$, $\bar{X} \longrightarrow N(\mu, \sigma/\sqrt{n})$.
  - $X_i \sim N(\mu,\sigma)$, $Z_i = \frac{X_i-\mu}{\sigma} \sim N(0,1)$ y $\sum Z_i \sim \chi^2_n$.
  - $X_i \sim N(\mu,\sigma)$, $T = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$.

- Teorema (Fisher) . $X_i \sim N(\mu,\sigma)$ entonces:
  $\bar{X} \perp S^2$, $\frac{(n-1)S}{\sigma^2} \rightarrow \chi^2_{n-1}$, $\frac{\bar{X}-\mu}{S/\sqrt{n}} \rightarrow t_{n-1}$.

- Distribución de la proporción muestral:
  $X_i \sim Ber(p)$, $X = \sum X_i \sim Bin(n,p)$; $p = \frac{X}{n}$.

Ya que $\mathbb{E}[p] = \frac{np}{n} = p$, $\operatorname{var}(p) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$

$$\left[\begin{aligned} \mathbb{E}[X^2] &= n\mathbb{E}[X_i^2] + 2\binom{n}{2}\mathbb{E}[X_iX_j] \\ &= np + n(n-1)p^2 \\ \mathbb{E}[X]^2 &= (np)^2 \qquad \therefore \operatorname{var}(X) = np - np^2 \\ &\qquad\qquad\qquad\qquad = np(1-p). \end{aligned}\right]$$

$\hookrightarrow$ Instead, $\operatorname{var}(X) = \frac{1}{n^2}\operatorname{var}(X_i) = n\frac{p(1-p)}{n^2} = \frac{p(1-p)}{n}$.

$\hookrightarrow$ by CLT $p \longrightarrow N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$.

- Distribución de la diferencia de proporciones:
  $X_i \sim Ber(p_X)$, $Y_i \sim Ber(p_Y)$ $\Rightarrow \hat{p}_X - \hat{p}_Y = \frac{X}{n_X} - \frac{Y}{n_Y}$
  $X = \sum X_i$ $\quad Y = \sum Y_i$

  $\therefore$ CLT says $\hat{p}_X - \hat{p}_Y \longrightarrow N\left(p_X - p_Y, \sqrt{\frac{p_X(1-p_X)}{n_X} + \frac{p_Y(1-p_Y)}{n_Y}}\right)$

⚠ Página 48/380 pdf tabla resumen.

CAPÍTULO 2 - Estimación puntual.

- Un estadístico muestral que quiere aproximar el valor de un parámetro se llama estimador. Este proceso se llama estimación. También se puede llevar a cabo una contrastación de hipótesis.
  ↳ $\theta(X_1, \dots, X_n)$: estimador
  $\theta(x_1, \dots, x_n)$: estimación puntual

- Error cuadrático medio del estimador $\hat{\theta}$ es $\mathrm{Ecm}(\hat{\theta}) = \mathbb{E}[(\hat{\theta}-\theta)^2] = \mathrm{var}(\hat{\theta}) + (\mathbb{E}[\hat{\theta}]-\theta)^2$
  $= \mathrm{var}(\hat{\theta}) + \underbrace{\mathrm{sesgo}}_{:= \text{bias!}}(\hat{\theta})^2$

- $\hat{\theta}$ es insesgado/centrado si $\mathrm{sesgo}(\hat{\theta}) = 0$.
  ej: $S^2 = \frac{1}{n-1}\sum(X_i - \bar{X})^2$, $\mathbb{E}(S^2) = \sigma^2$ ∴ centrado!

- $\hat{\theta}_0$ es insesgado y uniformemente de mínima varianza (UMVUE) si $\mathrm{var}(\hat{\theta}_0) \leq \mathrm{var}(\hat{\theta})$.

- Cota de Frechet-Cramér-Rao:
  $$\mathrm{var}(\hat{\theta}) \geq \frac{[1 + b'(\theta)]^2}{\mathbb{E}\left[\left(\frac{\partial \ln f(x_1, \dots, x_n; \theta)}{\partial \theta}\right)^2\right]}, \quad b(\theta) = \mathrm{sesgo}(\theta)$$
  $= I(\theta)$

- Un estimador es eficiente si es insesgado y tiene mínima varianza (i.e. ⟺ $\mathrm{var}(\hat{\theta})$ = cota FCR). La eficiencia es $= \frac{\text{cota FCR}}{\mathrm{var}(\hat{\theta})}$, eficiencia relativa de $\hat{\theta}_1, \hat{\theta}_2$ es $\frac{\mathrm{var}(\hat{\theta}_1)}{\mathrm{var}(\hat{\theta}_2)}$,

- $\{\hat{\theta}_n\}$ son consistentes si $\hat{\theta}_n \xrightarrow{\mathbb{P}} \theta$.
  Son consistentes en media cuadrática si $\hat{\theta}_n \xrightarrow{\mathcal{L}^2} \theta$.
  (in $\mathcal{L}^2$ ⟹ in prob).

- $T(X_1, \dots, X_n)$ es suficiente para el parámetro θ si $f_T(x_1, \dots, x_n \mid T = t)$ no depende de θ.

  T es suficiente si y solo si ∃ una decomposición
  $f(x_1, \dots, x_n; \theta) = g(T(x_1, \dots, x_n); \theta)\, h(x_1, \dots, x_n)$.

- Eficiente ⟹ suficiente.

CAPÍTULO 3 – Métodos de obtención de estimadores

- Método de los momentos:

  ① Supongamos que hay k parámetros $\theta_1, \dots, \theta_k$.
  Let $\alpha_i = \mathbb{E}[X^i] = \int x^i f(x; \theta_1, \dots, \theta_k)\,dx$.

  ② Let $a_1 = \frac{1}{n}\sum x_i$, $a_2 = \frac{1}{n}\sum x_i^2$, $\dots$, $a_k = \frac{1}{n}\sum x_i^k$.

  ③ Halla las solución a $\begin{cases} \alpha_1(\theta_1, \dots, \theta_k) = a_1 \\ \vdots \qquad\qquad \vdots \\ \alpha_k(\theta_1, \dots, \theta_k) = a_k \end{cases}$.

  La solución $\hat{\theta}_1, \dots, \hat{\theta}_k$ son los estimadores.

  ↳ Propiedades
  - Insesgadez
  - Consistentes.
  - Normalidad asintótica

  } Aún así no es un muy método (mejor MLE).

- Método de la máxima verosimilitud:

  Dejemos que $L(x; \theta) = f(x_1, \dots, x_n; \theta)$, lo que llamamos verosimilitud. Entonces,

  $$\hat{\theta}(x_1, \dots, x_n) = \underset{\theta \in \Omega}{\operatorname{argmax}}\, L(x_1, \dots, x_n; \theta).$$

  ↳ Estimador de máx. verosimilitud (EMV).

  Propiedades:
  - Consistentes
  - Asintóticamente insesgados
  - Si $\exists\, \hat{\theta}$ eficientes $\Rightarrow \hat{\theta}$ es EMV ($\Leftarrow$ no es cierto).
  - Eficiencia asintótica
  - $\hat{\theta} \longrightarrow N(\theta, \sqrt{var(\hat{\theta})})$

    donde $var(\hat{\theta}) = \dfrac{1}{n\mathbb{E}[I'(\theta)^2]}$ = cota de FCR.
  - Suficiente

CAPÍTULO 4

- Un intervalo de confianza $[\underline{\theta}(x_1,\dots,x_n), \overline{\theta}(x_1,\dots,x_n)]$ es un intervalo aleatorio. confianza del $100(1-\alpha)\%$.
  si $P[\underline{\theta}(x_1,\dots,x_n) \le \theta \le \overline{\theta}(x_1,\dots,x_n)] = 1-\alpha$.
  ↳ Intervalos bilaterales

- Métodos:

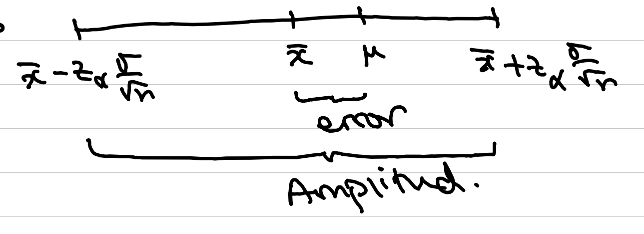
① método pivotal: escoger $T(x_1,\dots,x_n;\theta)$ que dependa de θ,
   ej: $X_i \sim N(\mu,\sigma) \Rightarrow \bar{X} = \frac{1}{n}\sum X_i \sim N(\mu, \frac{\sigma}{\sqrt{n}})$
   $\Rightarrow \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$.
   $var = \frac{1}{n^2}(\sigma^2+\dots+\sigma^2) = \sigma^2/n$
   so std $= \sigma/\sqrt{n}$.
   ...

② Método de Neyman.
   ?

- $N(\mu,\sigma^2)$, σ conocido $\Rightarrow \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \to N(0,1)$
  $N(\mu,\sigma^2)$, σ desconocido $\Rightarrow \dfrac{\bar{X}-\mu}{S/\sqrt{n}} \to t_{n-1}$
  $N(\mu,\sigma^2)$, μ desconocida $\to \dfrac{(n-1)S^2}{\sigma^2} \to \chi^2_{n-1}$.
  μ conocido $\Rightarrow \dfrac{\sum(X_i-\mu)^2}{\sigma^2} \to \chi^2_n$.
  ... similar para $X_i \sim N(\mu_X,\sigma_X^2)$, $Y_i \sim N(\mu_Y,\sigma_Y^2)$.

- Casos no normales:
  * Desigualdad de Chebychev: si $E[X_i]=\mu$, $var(X_i)=\sigma^2$ entonces
    $P(|\bar{x}-\mu| \ge k\frac{\sigma}{n}) \le \frac{1}{k^2}$
    $\Rightarrow P(|\bar{x}-\mu| \cancel{\ge} k) \le 1-\frac{\sigma^2}{nk^2}$ ...
  * muestras grandes (utilizar CLT):
    ① $\hat{\theta} \to N(\theta, \sqrt{var(\hat{\theta})})$, $var(\hat{\theta}) = \dfrac{1}{nE[I'(\theta)^2]}$
    $\therefore \dfrac{\hat{\theta}-\theta}{\sqrt{var(\hat{\theta})}} \to N(0,1)$ ... $I(\theta) = \ln f(x;\theta)$
    ② CLT: sabemos que en general
    $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \to N(0,1)$.

- Proporción (para muestras pequeñas):
  EMV: $\hat{p} = \frac{X}{n}$, $P(\hat{p}=x)$
  $= P(X=nx) = \binom{n}{nx} p^{nx}(1-p)^{n-nx}$
  for $x = 0, \frac{1}{n}, \dots, n$.
  Así que para encontrar C.I
  has de hacer manualment $P(\hat{p} \le p_0)$...
  hasta llegar a $P(\hat{p} \le p_0) \le \alpha/2$, $P(\hat{p} \ge p_1) \le \alpha/2$
  $\therefore CI = [p_0, p_1]$.
  (Para muestras grandes usa ①).

- 

CAPÍTULO 5

- Hipótesis paramétricas si $\mathcal{H}$: si θ .... .
  - ↳ Simple $\mathcal{H}: \theta = \theta_0$
  - compuesta $\mathcal{H}: \theta \neq \theta_0$ o $\theta > \theta_0$ o ...

- Hipótesis nula $\mathcal{H}_0$ se acepta provisionalmente.
  $\mathcal{H}_0: \theta \in \Omega_0 \subseteq \Omega$. Hipótesis alternativa,
  $\mathcal{H}_1: \theta \in \Omega_1 \subseteq \Omega$ tal que $\Omega_0 \cap \Omega_1 = \emptyset$
  $\Omega_0 \cup \Omega_1 = \Omega$.
  $\mathcal{H}_0$ es cierto si $\theta \in \Omega_0$, falsa cuando $\theta \in \Omega_0^c = \Omega_1$.

- El conjunto de observaciones muestrales tal que rechazamos $\mathcal{H}_0$ es la región crítica C.
  El complemento es la región de aceptación $\overline{C}$.
  $\therefore \mathbb{R}^n = C \cup \overline{C}$.

- Errores

	$\mathcal{H}_0$ cierta	$\mathcal{H}_0$ falsa
Aceptamos $\mathcal{H}_0$	✓	Error tipo II β
rechazamos $\mathcal{H}_0$	Error tipo I α	✓

- Talla o tamaño para el error de tipo I es
  $$\alpha = \max_{\theta \in \Omega_0} \mathbb{P}(\text{rechazo } \mathcal{H}_0 \mid \mathcal{H}_0 \text{ es cierta})$$
- Talla o tamaño para el error de tipo II
  $$\beta = \max_{\theta \in \Omega_1} \mathbb{P}(\text{Acepto } \mathcal{H}_0 \mid \mathcal{H}_0 \text{ es falsa}).$$
- Un test es ideal si $\alpha = \beta = 0$.

- Nótese que el nivel de confianza es precisamente $1-\alpha$ (aunque aquí ignoramos β).

- La potencia de un test es $= \mathbb{P}(\text{Rechazar } \mathcal{H}_0 \mid \theta)$
  $$\begin{cases} \cdot \text{ si } \theta \in \Omega_0, \text{ potencia} = \alpha(\theta) \\ \cdot \text{ si } \theta \in \Omega_1, \text{ potencia} = 1-\beta(\theta). \end{cases}$$
  $\therefore$ Queremos un test con una potencia grande en Ω_1 y pequeña en Ω_0.

Fases para realizar un contraste o test de hipótesis:
1. Formular $\mathcal{H}_0, \mathcal{H}_1$
2. Determinar un test estadístico
3. Seleccionar α
4. Determinar la región crítica
5. Seleccionar aleatoriamente la muestra y determinar el resultado.

CAPÍTULO 6

- ej: $X_i \sim N(\mu, \sigma^2)$, σ conocida, then

$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ por lo tanto si

$\mathcal{H}_0: \mu = \mu_0$ $\Rightarrow$ $\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$ en $\mathcal{H}_0$

$\mathcal{H}_1: \mu \neq \mu_0$

$\therefore -z_{\alpha/2} \leq \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} < z_{\alpha/2}$

región de aceptación

Si $\mathcal{H}_0: \mu \geq \mu_0$, $\mathcal{H}_1: \mu < \mu_0$ } Aceptación si: $-z_{\alpha/2} < \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$.

- Relación entre I.C y contraste de hipótesis:

Si nuestro I.C al nivel $100(1-\alpha)\%$ es $[\underline{\theta}, \bar{\theta}]$ quiere decir que $\mathbb{P}(\underline{\theta} \leq \theta \leq \bar{\theta}) = 1-\alpha$.
Si contrastamos la hipótesis $\mathcal{H}_0: \theta = \theta_0$, $\mathcal{H}_1: \theta \neq \theta_0$
entonces rechazamos $\mathcal{H}_0$ si $\theta \notin [\underline{\theta}, \bar{\theta}]$.

- $N(\mu, \sigma^2)$, σ desconocido ... $t = \frac{\bar{x}-\mu_0}{S/\sqrt{n}}$ $\rightarrow t_{n-1}$ $n<30$, $\rightarrow N(0,1)$ $n \geq 30$

donde $S = \frac{1}{n-1}\sum(x_i-\bar{x})^2$ ...

$\boxed{\text{Tipo de t-test}}$ — def: any test where the t-distn is used.

- $N(\mu, \sigma^2)$ queremos contrastar sobre σ, $\chi^2 := \frac{1}{\sigma_0^2}\sum(x_i-\mu)^2 \sim \chi^2_n$ ....

Si μ es desconocido $\chi^2 = \frac{1}{\sigma_0^2}\sum(x_i-\bar{x})^2 \sim \chi^2_{n-1}$.

⋮ Lo mismo para el resto — memorizar tabla?

- El p-valor es el menor nivel de significación observado para que la hipótesis nula sea rechazada.

ie. $\mathbb{P}(Z > z)$ or $\mathbb{P}(T > t)$ for a test test...

CAPÍTULO 7

- Test no paramétricos — menos potentes.

- **Contraste de bondad de ajuste**:

  * contraste χ^2 de Pearson de bondad de ajuste

  X variable aleatoria, distribución $F_0(x)$, $S_1, \dots, S_k$ partición de subconjuntos de X,
  $p_i = \mathbb{P}(X \in S_i)$, $\sum p_i = 1$.

  $\mathcal{H}_0$: muestra procede de $F_0(x)$
  $\mathcal{H}_1$: " no "

  let $O_i = \#$ observaciones en $X \in S_i = n_i$
  $\therefore n_1 + \dots + n_k = n$.

  $$\therefore \mathbb{P}(n_1, \dots, n_k) = \frac{n!}{n_1! \cdots n_k!} p_1^{n_1} \cdots p_k^{n_k}$$

  y $N_i \sim B(n, p_i)$

  Si es correcto esperamos que $n_i \approx \mathbb{E}[N_i] = np_i$

  $$\therefore \chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i} \sim \chi^2_{k-1}.$$

  $\hookrightarrow \chi^2 =$ estadístico de bondad de ajuste.

  * Si $\mathcal{H}_0 : F \overset{def}{=} F_0(x; \theta_1, \dots, \theta_k)$, $\mathcal{H}_1 : F(x) \neq F_0$
  y $\hat{\theta}_i$ so los EMV entonces

  $$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_i(\hat{\theta}_1, \dots, \hat{\theta}_k))^2}{np_i(\hat{\theta}_1, \dots, \hat{\theta}_k)} \longrightarrow \chi^2_{k-h-1}.$$

  * Contraste de Kolmogorov - Smirnov:

  let $F_n(x) = \dfrac{N(x)}{n}$, $N(x) = \#$ observaciones $\le x$.

  let $D_n = \max_{x \in \mathbb{R}} | F_0(x) - F_n(x) |$

  El test para $\mathcal{H}_0 : F(x) = F_0(x)$ utiliza
  $\mathbb{P}(D_n > D_\alpha \mid \mathcal{H}_0) = \alpha$.
  se obtiene en una tabla

  * Contraste de normalidad de Lilliefors.

- **Contrastes de aleatoriedad**

  $\hookrightarrow$ Testeamos si X_i son iid.

  * Contraste de rachas de Wald — Wolfowitz.

  $\hookrightarrow$ eg. HTHHHT — parece aleatorio; HHHHHTTTT — no parece.

  Análisis: let $R = \#$ de rachas $= R_1 + R_2$ ($R_1 = \#$ rachas H, $R_2 = \#$ rachas T). Si observamos n_1 Hs y n_2 Ts, $n = n_1 + n_2$. Entonces,

  $$\mathbb{P}(R = r) = \frac{2\binom{n_1 - 1}{r/2 - 1}\binom{n_2 - 1}{r/2 - 1}}{\binom{n}{n_1}}$$

  r par. — Escoges donde empezar cada racha por H y T. — colocar las H es suficiente

  Existen tablas por lo que $\mathbb{P}(R \le k_{\alpha/2}) \le \alpha/2$
  y pues $[0, k'_{\alpha/2}]$, $[k_{\alpha/2}, n]$ es la región crítica.

  Para $R \to \infty$, $R \longrightarrow N(\mathbb{E}[R], \sqrt{\operatorname{var} R})$, con $\mathbb{E}[R] = \dfrac{2n_1 n_2}{n} + 1$ y $\operatorname{var} R = \dfrac{2n_1 n_2(2n_1 n_2 - n)}{n^2(n-1)}$

  * Si hay variaciones cíclicas no funciona.

- **Contraste de localización**:

  * muestra de una distⁿ $F(x)$, $0 < p < 1$ y $C_p(F)$ es el cuantil de orden p de $F(x)$. Hipótesis:

  $\mathcal{H}_0 : C_p(F) = k_0$ $\rightsquigarrow$ ie. $\mathbb{P}(X \le k_0) = p$
  $\mathcal{H}_1 : C_p(F) \neq k_0$.

  let $S = \#\{X_i > k_0\}$ so $\mathbb{P}(S = s) = \binom{n}{s} q^s p^{n-s}$.

  $\therefore$ Utilizar $\mathbb{P}(S \le k_{\alpha/2}) \approx \alpha/2$ ...

  * Similar si conocemos la mediana en distribuciones simétricas (test de rangos-signos de Wilcoxon).

Población	Queremos	Estadístico + dist.
$N(\mu,\sigma^2)$ σ conocido	μ	$z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
$N(\mu,\sigma^2)$ σ descon.	μ	$t = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$ donde $S^2 = \frac{1}{n-1}\sum(X_i-\bar{X})^2 \to \sigma$
$N(\mu,\sigma^2)$ μ conocido	σ^2	$\chi^2 = \frac{\sum(x_i-\mu)^2}{\sigma^2} \sim \chi^2_n$
$N(\mu,\sigma^2)$ μ descon.	σ^2	$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \frac{\sum(X_i-\bar{X})^2}{\sigma^2} \sim \chi^2_{n-1}$
$N(\mu_x,\sigma_x^2)$ $N(\mu_y,\sigma_y^2)$ σ_x,σ_y con.	$\mu_x-\mu_y$	$z = \frac{\bar{X}-\bar{Y}-(\mu_x-\mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x}+\frac{\sigma_y^2}{n_y}}} \sim N(0,1)$
$N(\mu_x,\sigma_x^2)$ $N(\mu_y,\sigma_y^2)$ σ_x,σ_y desc. $\sigma_x=\sigma_y$	$\mu_x-\mu_y$	$t = \frac{\bar{X}-\bar{Y}-(\mu_x-\mu_y)}{s'\sqrt{1/n_x+1/n_y}} \sim t_{n_x+n_y-2}$. $(s')^2 = \frac{(n_x-1)s_x^2+(n_y-1)^2 s_y}{n_x-1+n_y-1}$
If $\sigma_x \neq \sigma_y$	$\mu_x-\mu_y$	$t = \frac{\bar{X}-\bar{Y}-(\mu_x-\mu_y)}{\sqrt{S_x^2/n_x+S_y^2/n_y}} \sim t_v$ $v = \ldots ?$
$N(\mu_x,\sigma_x^2)$ $N(\mu_y,\sigma_y^2)$ μ_x,μ_y con.	test $\sigma_x^2=\sigma_y^2$	$F = \frac{1/n_x \sum(x_i-\mu_x)^2}{1/n_y \sum(y_i-\mu_y)^2} \sim F_{n_x,n_y}$.
$B(1,p)$	Test $p=p_0$	$\hat{p}=\frac{x}{n}$, $z = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim N(0,1)$.