,

Clase 3- Diferentes convergencias y Teorema Central del Límite

Probabilidades

Carrera de Especialización en Estadística,
Facultad de Ciencias Exactas y Naturales, UBA

2023

Sobre las encuestas

1. Me parece importante que las clases no se pasen del horario estipulado, ya que es injusto para quienes, por algún motivo, se tienen que ir a la hora acordada.
2. Me gusta que haya rigor matemático y que tenga un ritmo ágil, aunque a veces se me dificulta tomar apuntes.
3. Siento que las explicaciones teóricas se temrinan yendo mucho a demostraciones que no termino de comprender

¿Qué vamos a ver hoy?

- Diferentes nociones de convergencia.
- Demostrar desigualdades pendientes y Ley de los grande números
- Hablar de convergencia en distribución
- Teorema Central del Límite.

Desigualdades

- Markov: $X \geq 0$, entonces para todo $\delta > 0$ vale que

$$\mathbb{P}(X \geq \delta) \leq \frac{\mathbb{E}[X]}{\delta}$$

- Tchebycheff : Sea W una v.a. y $\varepsilon > 0$

$$\mathbb{P}\left(|W - \mathbb{E}[W]| \geq \varepsilon\right) \leq \frac{V[W]}{\varepsilon^2}$$

Demostración:

El mundo normal vs. Desigualdad Tchebycheff

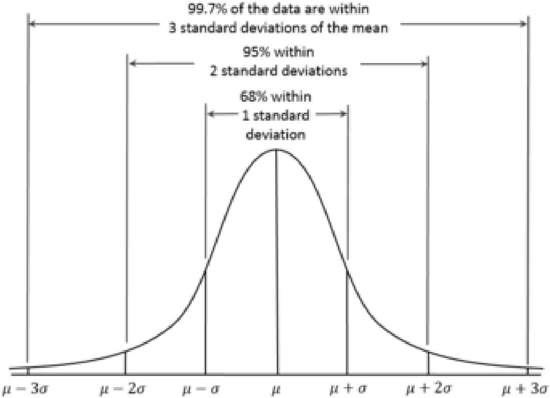
- Probabilidades exactas en el mundo normal

$$\mathbb{P}\left(|W - \mu_W| > \varepsilon\right) = 2\left(1 - \phi\left(\frac{\varepsilon}{\sigma_W}\right)\right)$$

- Cotas superiores universales (asumiendo $V(W)$ finita)

$$\mathbb{P}\left(|W - \mu_W| > \varepsilon\right) \leq \frac{\sigma_W^2}{\varepsilon^2}$$

Regla Normal - $W \sim \mathcal{N}(\mu, \sigma^2)$



$$
\begin{aligned}
\mathbb{P}\left(|W - \mu| \geq \sigma\right) &= 0{,}32 \;, \\
\mathbb{P}\left(|W - \mu| \geq 2\sigma\right) &= 0{,}05 \;, \\
\mathbb{P}\left(|W - \mu| \geq 3\sigma\right) &= 0{,}03
\end{aligned}
$$

Desigualdad de Tchebycheff

Sea W una v.a. con media $\mathbb{E}(W) = \mu$ y $V(W) = \sigma^2$. Luego, $\varepsilon > 0$

$$\mathbb{P}\left(|W - \mu| \geq \varepsilon\right) \leq \frac{\sigma^2}{\varepsilon^2}$$

Tomando $\varepsilon = k\sigma$,

$$\mathbb{P}\left(|W - \mu| \geq k\sigma\right) \leq \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2}$$

Desigualdad de Tchebycheff

Sea W una v.a. con media $\mathbb{E}(W) = \mu$ y $V(W) = \sigma^2$. Luego, $\varepsilon > 0$

$$\mathbb{P}\left(|W - \mu| \geq \varepsilon\right) \leq \frac{\sigma^2}{\varepsilon^2}$$

Tomando $\varepsilon = k\sigma$,

$$\mathbb{P}\left(|W - \mu| \geq k\sigma\right) \leq \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2}$$

En particular, para $k = 1, 2, 3$, obtenemos las las siguientes cotas:

$$\begin{aligned}
\mathbb{P}\left(|W - \mu| \geq \ \sigma\right) &\leq \ 1 \ , \\
\mathbb{P}\left(|W - \mu| \geq \ 2\sigma\right) &\leq \ \frac{1}{4} \ , \\
\mathbb{P}\left(|W - \mu| \geq \ 3\sigma\right) &\leq \ \frac{1}{9}
\end{aligned}$$

El mundo normal vs. Desigualdad Tchebycheff

- Probabilidades exactas en el mundo normal

$$
\begin{aligned}
\mathbb{P}\left(|W-\mu| \geq \sigma\right) &= 0{,}32 \, , \\
\mathbb{P}\left(|W-\mu| \geq 2\sigma\right) &= 0{,}05 \, , \\
\mathbb{P}\left(|W-\mu| \geq 3\sigma\right) &= 0{,}03
\end{aligned}
$$

- Cotas superiores universales (asumiendo $V(W)$ finita)

$$
\begin{aligned}
\mathbb{P}\left(|W-\mu| \geq \sigma\right) &\leq 1 \, , \\
\mathbb{P}\left(|W-\mu| \geq 2\sigma\right) &\leq \frac{1}{4} \, , \\
\mathbb{P}\left(|W-\mu| \geq 3\sigma\right) &\leq \frac{1}{9}
\end{aligned}
$$

Desigualdad de Tchebycheff aplicada a promedios

- $(X_i)_{i\geq 1}$ iid, $\mathbb{E}(X_i)=\mu$, $V(X_i)=\sigma^2$.
- Promedio :

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \sim ?$$

- Esperanza y Varianza del promedio

$$\mu_{\overline{X}_n} = \mathbb{E}(\overline{X}_n) = \mu \ , \quad \sigma^2_{\overline{X}_n} = V(\overline{X}_n) = \frac{\sigma^2}{n} \ .$$

- Tchebycheff dice:

$$\mathbb{P}\left(|\overline{X}_n - \mu_{\overline{X}_n}| \geq \varepsilon\right) \ \leq \ \frac{\sigma^2_{\overline{X}_n}}{\varepsilon^2}$$

$$\mathbb{P}\left(|\overline{X}_n - \mu| \geq \varepsilon\right) \ \leq \ \frac{\sigma^2/n}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$$

Teorema: Ley de los Grandes Números (LGN)

Sean $(X_i)_{i \geq 1}$ i.i.d., con $\mathbb{E}(X_i) = \mu$ y $V(X_i) = \sigma^2$, para todo i.
Entonces, el promedio converge a μ en probabilidad:

es decir para todo $\varepsilon > 0$ vale que

$$\lim_{n \to \infty} \mathbb{P}\left(|\overline{X}_n - \mu| > \varepsilon\right) = 0$$

$$\overline{X}_n \xrightarrow{p} \mu$$

Demostración de la Ley de los Grandes Números

Demostración: Tchebycheff prueba que

$$\mathbb{P}\left(|W - \mathbb{E}[W]| \geq \varepsilon\right) \leq \frac{V[W]}{\varepsilon^2}$$

Vamos a invocar la Desigualdad de Tchebicheff, pero con $W = \overline{X}_n$. En tal caso,

$$\mu_{\overline{X}_n} = \mathbb{E}(\overline{X}_n) = \mu$$

$$\sigma^2_{\overline{X}_n} = V(\overline{X}_n) = \frac{\sigma^2}{n}\,.$$

Tenemos entonces que

$$\mathbb{P}\left(|\bar{X}_n - \mu| > \varepsilon\right) = \mathbb{P}\left(|\bar{X}_n - \mu_{\overline{X}_n}| > \varepsilon\right) \leq \frac{V(\overline{X}_n)}{\varepsilon^2} = \frac{1}{\varepsilon^2}\frac{\sigma^2}{n} \to 0$$

Muestra (aleatoria simple)- variable i.i.d

Diremos que $X_1, \ldots, X_n$ son una muestra si son variables aleatorias independientes, idénticamente distribuídas.

$$X_1, \ldots, X_n \text{ , i.i.d.}$$

En tal caso, $X_i \sim F$ para todo i, y por consiguiente,

$$\text{mongo}(X_i) = \text{mongo}(X_j) = \text{mongo}(X_1).$$

LGN - más generalmente

$$\frac{1}{n}\sum_{i=1}^{n}\texttt{mongo}_i \to \mathbb{E}(\texttt{mongo}) \quad \text{en probabilidad}$$

Utilizamos la Ley de los Grandes Números para Estimación

Estimación de la Esperanza

- $(X_i)_{i \geq 1}$, i.i.d. $X_i \sim F$.
- $\mathbb{E}(X_i) = \mu$, $V(X_i) = \sigma^2$
- Por la ley de los grandes números

$$\overline{X}_n \to \mathbb{E}(X_1) = \mu \quad \text{en probabilidad}$$

- En tal caso, definimos

$$\widehat{\mu} = \widehat{\mu}_n = \overline{X}_n$$

.

- Propiedades de $\widehat{\mu}_n$:
  - $\mathbb{E}(\widehat{\mu}_n) = \mu$
  - $\widehat{\mu}_n \to \mu$ en probabilidad.

Variables Bernoulli

- $X_i \sim \mathcal{B}(1, p)$, i.i.d.
- $\mathbb{E}(X_i) = p$, $V(X_i) = p(1-p)$
- Por la ley de los grandes números

$$\overline{X}_n \to \mathbb{E}(X_1) = p \quad \text{en probabilidad}$$

- En tal caso, definimos

$$\widehat{p} = \widehat{p}_n = \overline{X}_n$$

.

- Propiedades de $\widehat{p}_n$:
  - $\mathbb{E}(\widehat{p}_n) = p$
  - $\widehat{p}_n \to p$ en probabilidad.

Estimación de Probabilidades - Ejemplo

- $(X_i)_{i\geq 1}$, i.i.d. $X_i \sim F$.
- Nos interesa estimar $F(3) = \mathbb{P}(X \leq 3)$.
- Definimos
$$Y_i = I_{X_i \leq 3} = I_{(-\infty,3]}(X_i)$$
- $Y_i \sim \mathcal{B}(1,p)$, con
$$p = \mathbb{E}(Y_1) = \mathbb{P}(Y_1 = 1) = \mathbb{P}(X_1 \leq 3) = F(3)$$
- Por la ley de los grandes números (caso Bernoulli)
$$\overline{Y}_n \to \mathbb{E}(Y_1) = F(3) \quad \text{en probabilidad}$$
- En tal caso, definimos
$$\widehat{F}(3) = \widehat{F}_n(3) = \overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} I_{X_i \leq 3}$$

.

Estimación de Probabilidades.

- $(X_i)_{i\geq 1}$, i.i.d. $X_i \sim F$.
- Nos interesa estimar $\mathbb{P}(X \in A)$.
- Definimos
$$Y_i = I_{X_i \in A} = I_A(X_i)$$
- $Y_i \sim \mathcal{B}(1, p)$, con
$$p = \mathbb{E}(Y_1) = \mathbb{P}(Y_1 = 1) = \mathbb{P}(X_1 \in A)$$
- Por la ley de los grandes números (caso Bernoulli)
$$\overline{Y}_n \to \mathbb{E}(Y_1) = \mathbb{P}(X_1 \in A) \quad \text{en probabilidad}$$
- Es decir,
$$\frac{1}{n}\sum_{i=1}^{n} I_{X_i \in A} \longrightarrow \mathbb{P}(X_1 \in A)$$

DEMOSTRAMOS QUE LA FRECUENCIA RELATIVA CONVERGE A LA PROBABILIDAD. 👏👏

Estimación de varianza.

- $(X_i)_{i\geq 1}$, i.i.d. $X_i \sim F$.
- Nos interesa estimar $\sigma^2 = V(X) = \mathbb{E}(X^2) - \{\mathbb{E}(X)\}^2$.
- ¿Qué hacemos?

LGN y Propiedades de la convergencia en probabilidad

- Ley de los grandes Números: $\texttt{mongo}_1, \texttt{mongo}_2, \ldots$ iid, $\texttt{mongo}_i \sim \texttt{mongo}$

$$\frac{1}{n}\sum_{i=1}^{n} \texttt{mongo}_i \quad \xrightarrow{p} \quad \mathbb{E}(\texttt{mongo})$$

LGN y Propiedades de la convergencia en probabilidad

- Ley de los grandes Números: $\texttt{mongo}_1, \texttt{mongo}_2, \ldots$ iid, $\texttt{mongo}_i \sim \texttt{mongo}$

$$\frac{1}{n}\sum_{i=1}^{n}\texttt{mongo}_i \quad \xrightarrow{p} \quad \mathbb{E}(\texttt{mongo})$$

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{p} \mathbb{E}(X^2)$$

- Si $W_n \xrightarrow{p} W$ y g continua, entonces $g(W_n) \xrightarrow{p} g(W)$.

$$\left\{\overline{X}_n\right\}^2 \xrightarrow{p} \{\mathbb{E}(X)\}^2$$

- Si $W_n \xrightarrow{p} W$ y $U_n \xrightarrow{p} U$, entonces $W_n \pm U_n \xrightarrow{p} W \pm U$.

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \{\overline{X}_n\}^2 \xrightarrow{p} \mathbb{E}(X^2) - \mu^2 = Var(X)$$

Propiedades de la convergencia en probabilidad

- Si $W_n \xrightarrow{p} W$ y $U_n \xrightarrow{p} U$, entonces $W_n U_n \xrightarrow{p} WU$.
- Si $W_n \xrightarrow{p} W$ y $U_n \xrightarrow{p} U$ y $\mathbb{P}(U = 0) = 0$, entonces $W_n / U_n \xrightarrow{p} W/U$.
- Si $W_n \xrightarrow{p} W$ y $U_n \xrightarrow{p} U$ y g continua, entonces $g(W_n, U_n) \xrightarrow{p} g(W, U)$.

Volvamos a mirar como se distribuyen las *cosas*

Convergencia en distribución

Definition

Sea $W_1, W_2, \ldots$ una sucesión de variables aleatorias y sea W otra variable aleatoria. Denotemos con F_n y F a la función de distribución acumulada de W_n y W, respectivamente: $W_n \sim F_n$, $W \sim F$. Diremos que W_n converge en distribución a W si

$$\lim_{n\to\infty} F_n(t) = F(t), \quad \text{para todo } t \text{ punto de continuidad de } F$$

Notación: $W_n \xrightarrow{D} W$.

Teorema Central del límite

Sea $X_1, X_2, \ldots$ una sucesión de variables aleatorias iid definidas en $(\Omega, \mathcal{A}, P)$, con $\mathbb{E}(X_i) = \mu$ y $V(X_i) = \sigma^2$, Entonces,

$$\frac{\overline{X}_n - \mu}{\sqrt{\mathbb{V}(\overline{X}_n)}} \xrightarrow{D} Z, \quad Z \sim \mathcal{N}(0,1).$$

Versiones equivalentes

$$\frac{\overline{X}_n - \mu}{\sqrt{\mathbb{V}(\overline{X}_n)}} \xrightarrow{D} Z, \quad Z \sim \mathcal{N}(0,1).$$

$$\frac{\overline{X}_n - \mu}{\sqrt{\sigma^2/n}} \xrightarrow{D} Z, \quad Z \sim \mathcal{N}(0,1).$$

$$\frac{\sqrt{n}\{\overline{X}_n - \mu\}}{\sqrt{\sigma^2}} \xrightarrow{D} Z, \quad Z \sim \mathcal{N}(0,1).$$

Teorema de Slutzky

1. Si $W_n \xrightarrow{D} W$ y $U_n \xrightarrow{p} c$, con $c \in \mathbb{R}$, entonces

$$W_n + U_n \xrightarrow{D} c + U$$

2. Si $W_n \xrightarrow{D} W$ y $U_n \xrightarrow{p} c$, con $c \in \mathbb{R}$, entonces

$$W_n U_n \xrightarrow{D} cU$$

Teorema de Slutzky

1. Si $W_n \xrightarrow{D} W$ y $U_n \xrightarrow{p} c$, con $c \in \mathbb{R}$, entonces
$$W_n + U_n \xrightarrow{D} c + U$$

2. Si $W_n \xrightarrow{D} W$ y $U_n \xrightarrow{p} c$, con $c \in \mathbb{R}$, entonces
$$W_n U_n \xrightarrow{D} cU$$

¿Cómo se usan estas cosas?

Método Delta

Supongamos que

$$\frac{\sqrt{n}W_n - \mu}{\sigma} \xrightarrow{D} \mathcal{N}(0,1)$$

y que g es una función diferenciable tal que $g'(\mu) \neq 0$. Entonces

$$\frac{\sqrt{n}g(W_n) - g(\mu)}{|g'(\mu)|\sigma} \xrightarrow{D} \mathcal{N}(0,1)$$

Es decir,

si $\quad W_n \approx \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right) \quad$ entonces $\quad g(W_n) \approx \mathcal{N}\left(g(\mu), \frac{(g'(\mu))^2\sigma^2}{n}\right)$

Vueltas y vueltas y vueltas...

Suma de variables Normales

- X_i i.i.d., $X_i \sim \mathcal{N}(\mu, \sigma^2)$
- Suma de normales es normal:

$$\sum_{i=1}^{n} X_i \sim \mathcal{N}(n\mu, n\sigma^2)$$

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n\sigma^2}} \sim \mathcal{N}(0,1)$$

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{n} X_i \leq u\right) &= \mathbb{P}\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n\sigma^2}} \leq \frac{u - n\mu}{\sqrt{n\sigma^2}}\right) \\
&= \phi\left((u - n\mu)/\sqrt{n\sigma^2}\right) \\
&= \texttt{pnorm}\left((u - n\mu)/\sqrt{n\sigma^2}\right)
\end{aligned}$$

Promedio de variables Normales

- X_i i.i.d., $X_i \sim \mathcal{N}(\mu, \sigma^2)$

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \sim \mathcal{N}\left(\mathbb{E}(\overline{X}_n), V(\overline{X}_n)\right) \ , \quad \overline{X}_n \sim \mathcal{N}\left(\mu, \sigma^2/n\right) \ ,$$

$$\frac{\overline{X}_n - \mu}{\sqrt{V(\overline{X}_n)}} \sim \mathcal{N}(0,1) \ , \ \frac{\overline{X}_n - \mu}{\sqrt{\sigma^2/n}} \sim \mathcal{N}(0,1) \ ,$$

$$\begin{aligned} \mathbb{P}\left(\overline{X}_n \leq u\right) &= \mathbb{P}\left(\frac{\overline{X}_n - \mu}{\sqrt{\sigma^2/n}} \leq \frac{u-\mu}{\sqrt{\sigma^2/n}}\right) \\ &= \phi\left(\sqrt{n}(u-\mu)/\sqrt{\sigma^2}\right) = \texttt{pnorm}\left(\sqrt{n}(u-\mu)/\sqrt{\sigma^2}\right) \end{aligned}$$

Teorema Central del Límite (TCL): versión suma

- Sean $(X_i)_{i\geq 1}$ v.a. i.i.d. con $\mathbb{E}(X) = \mu$ y $V(X) = \sigma^2$.
- ¿Qué sabemos de la suma ?

$$\sum_{i=1}^{n} X_i \ , \quad \mathbb{E}(\sum_{i=1}^{n} X_i) = n\mu, \quad V(\sum_{i=1}^{n} X_i) = n\sigma^2$$

TEOREMA: La distribución de la suma $\sum_{i=1}^{n} X_i$ SE PARECE a la de una normal, con su esperanza $\left(\mathbb{E}\left(\sum_{i=1}^{n} X_i\right)\right)$ y su varianza $\left(V(\sum_{i=1}^{n} X_i)\right)$

TCL: La suma tiene distribución Aproximadamente Normal

- X_i i.i.d., $\mathbb{E}(X_i) = \mu$, $V(X_i) = \sigma^2$

$$\sum_{i=1}^{n} X_i \overset{a}{\approx} \mathcal{N}\left(\mathbb{E}(\sum_{i=1}^{n} X_i), V(\sum_{i=1}^{n} X_i)\right) \ , \quad \sum_{i=1}^{n} X_i \overset{a}{\approx} \mathcal{N}\left(n\mu, n\sigma^2\right) \ ,$$

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n\sigma^2}} \overset{a}{\approx} \mathcal{N}(0,1) \ .$$

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^{n} X_i \leq t\right) &= \mathbb{P}\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n\sigma^2}} \leq \frac{t - n\mu}{\sqrt{n\sigma^2}}\right) \\ &\approx \texttt{pnorm}\left((t - n\mu)/\sqrt{n\sigma^2}\right) \end{aligned}$$

Teorema Central del Límite (TCL) - Suma:

Sean $(X_i)_{i\geq 1}$ v.a. i.i.d. $X \sim X_i$, con $\mathbb{E}(X) = \mu$ y $V(X) = \sigma^2$, entonces tenemos que

$$\mathbb{P}\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n\sigma^2}} \leq x\right) \underset{n\to\infty}{\longrightarrow} \phi(x) = \texttt{pnorm(x)}\ , \quad \forall x \in \mathbb{R}\ ,$$

TCL - Versión Promedio

- X_i i.i.d., $X \sim X_i$. $\mathbb{E}(X) = \mu$, $V(X) = \sigma^2$

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \ , \quad \mathbb{E}(\overline{X}_n) = \mu \ , \quad V(\overline{X}_n) = \sigma^2/n$$

TEOREMA: La distribución del promedio $\overline{X}_n$ SE PARECE a la de una normal, con su esperanza $(\mathbb{E}\left(\overline{X}_n\right))$ y su varianza $(V(\overline{X}_n))$

TCL: El promedio tiene distribución Aproximadamente Normal

- X_i i.i.d., $\mathbb{E}(X_i) = \mu$, $V(X_i) = \sigma^2$

$$\overline{X}_n \overset{a}{\approx} \mathcal{N}\left(\mathbb{E}(\overline{X}_n), V(\overline{X}_n)\right) \;, \quad \overline{X}_n \overset{a}{\approx} \mathcal{N}\left(\mu, \sigma^2/n\right) \;,$$

$$\frac{\overline{X}_n - \mu}{\sqrt{V(\overline{X}_n)}} \overset{a}{\approx} \mathcal{N}(0,1) \;, \; \frac{\overline{X}_n - \mu}{\sqrt{\sigma^2/n}} \overset{a}{\approx} \mathcal{N}(0,1) \;,$$

$$\mathbb{P}\left(\overline{X}_n \leq x\right) \approx \phi\left((x-\mu)/\sqrt{\sigma^2/n}\right)$$

$$\mathbb{P}\left(\overline{X}_n \leq x\right) \approx \texttt{pnorm}\left((x-\mu)/\sqrt{\sigma^2/n}\right)$$

Teorema Central del Límite (TCL) - Promedio:

Sean $(X_i)_{i \geq 1}$ v.a. i.i.d. $X \sim X_i$, con $\mathbb{E}(X) = \mu$ y $V(X) = \sigma^2$, entonces tenemos que

$$\mathbb{P}\left(\frac{\overline{X}_n - \mathbb{E}(\overline{X}_n)}{\sqrt{Var(\overline{X}_n)}} \leq x \right) \underset{n \to \infty}{\longrightarrow} \phi(x) = \texttt{pnorm(x)} \ , \quad \forall x \in \mathbb{R} \ ,$$

o equivalentemente

$$\mathbb{P}\left(\frac{\overline{X}_n - \mu}{\sqrt{\sigma^2 / n}} \leq x \right) \underset{n \to \infty}{\longrightarrow} \phi(x) = \texttt{pnorm(x)} \ , \quad \forall x \in \mathbb{R} \ .$$