

El método Bootstrap

26 de febrero de 2025

1. Introducción

A continuación, se presentan las ideas fundamentales del Bootstrap clásico, introducido por Efron (1979). Para la elaboración de estas notas, se toma como referencia principal el libro de Wasserman (2004).

Consideremos una muestra aleatoria Y_1, Y_2, \dots, Y_n proveniente de una variable aleatoria $Y \sim F_Y(y)$, donde asumimos que F_Y representa la distribución subyacente de un fenómeno aleatorio que queremos estudiar. Nuestro objetivo es estimar un funcional estadístico de la forma:

$$\theta = T(F_Y) = \int g(y) dF_Y(y).$$

Dado que la distribución F_Y es desconocida, podemos aproximarla mediante la función de distribución empírica:

$$\hat{F}_n(y) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{Y_i \leq y\}}. \quad (1)$$

Esta aproximación nos permite obtener de manera natural el estimador plug-in para θ :

$$\hat{\theta}_n = h(Y_1, Y_2, \dots, Y_n) = T(\hat{F}_n) = \frac{1}{n} \sum_{i=1}^n g(Y_i), \quad (2)$$

en donde estamos haciendo explícito que $\hat{\theta}_n$ se puede escribir como una función h de la m.a. Y_1, Y_2, \dots, Y_n y que conocemos esa función. El caso que se muestra en la expresión (2) es el caso sencillo, recordemos que también podríamos tener $\theta = \kappa(T_1(F_Y), T_2(F_Y), \dots, T_k(F_Y))$ pero se procedería de manera similar usando el estimador plug-in para cada $T_j(F_Y)$ y luego sustituyendo en la función κ .

Sin embargo, en inferencia estadística no solo nos interesa una estimación puntual de θ , sino también una medida de incertidumbre asociada, como $V(\hat{\theta}_n)$, o aun mejor un intervalo de confianza para θ . Para ello, necesitamos caracterizar la distribución muestral de $\hat{\theta}_n$. El método Bootstrap surge precisamente como una herramienta para aproximar esta distribución sin necesidad de suposiciones fuertes sobre F_Y . En las siguientes secciones, exploraremos su justificación mediante diversos escenarios hipotéticos hasta llegar a la formulación del algoritmo Bootstrap.

2. Re-muestreo Bootstrap

Obviamente, no conocemos $F_Y(y)$, pero hemos visto que $\hat{F}_n(y)$ es un excelente estimador. Si sustituimos la función de distribución acumulada empírica y ponemos unos asteriscos que indican que son muestras Bootstrap¹, obtenemos el siguiente algoritmo:

$$\begin{array}{rcl} Y_1^{*1}, Y_2^{*1}, \dots, Y_n^{*1} \stackrel{\text{m.a.}}{\sim} \hat{F}_n(y), & \rightarrow & \hat{\theta}_n^{*1} = h(Y_1^{*1}, Y_2^{*1}, \dots, Y_n^{*1}), \\ Y_1^{*2}, Y_2^{*2}, \dots, Y_n^{*2} \stackrel{\text{m.a.}}{\sim} \hat{F}_n(y), & \rightarrow & \hat{\theta}_n^{*2} = h(Y_1^{*2}, Y_2^{*2}, \dots, Y_n^{*2}), \\ \vdots & & \vdots, \\ Y_1^{*B}, Y_2^{*B}, \dots, Y_n^{*B} \stackrel{\text{m.a.}}{\sim} \hat{F}_n(y), & \rightarrow & \hat{\theta}_n^{*B} = h(Y_1^{*B}, Y_2^{*B}, \dots, Y_n^{*B}). \end{array}$$

Por lo tanto, podemos usar de nuevo la expresión (5) para estimar $V(\hat{\theta}_n)$, pero ahora insertando $\hat{\theta}_n^{*1}, \hat{\theta}_n^{*2}, \dots, \hat{\theta}_n^{*B}$. Entonces, tendríamos

$$V_{\text{Boot}}(\hat{\theta}_n) = \frac{1}{B} \sum_{j=1}^B (\hat{\theta}_n^{*j} - \bar{\hat{\theta}}_n^*)^2, \quad (5)$$

donde

$$\bar{\hat{\theta}}_n^* = \frac{1}{B} \sum_{j=1}^B \hat{\theta}_n^{*j}.$$

2.1. ¿Cómo generamos m.a. de $Y^* \sim \hat{F}_n(y)$?

La función de distribución acumulada empírica define una variable aleatoria discreta Y^* . Para cualquier variable aleatoria discreta W , la probabilidad de que tome un valor específico w está dada por:

$$P(W = w) = F_W(w) - \lim_{x \rightarrow w^-} F_W(x).$$

Podemos formalizar la propiedad fundamental de la función de distribución empírica en la siguiente proposición.

Proposición 1. *Sea $(Y_1, Y_2, \dots, Y_n) = (y_1, y_2, \dots, y_n)$ una muestra observada y sea $\hat{F}_n(y)$ su función de distribución acumulada empírica, definida como en la expresión (1). Entonces, la función de distribución empírica induce una variable aleatoria discreta Y^* que asigna una probabilidad de $\frac{1}{n}$ a cada observación y_i , es decir:*

$$P(Y^* = y_i) = \frac{1}{n}, \quad \text{para } i = 1, 2, \dots, n.$$

Demostración. La función de distribución empírica es una función de saltos con discontinuidades en los puntos de la muestra observada y_1, y_2, \dots, y_n , la probabilidad de que Y^* tome un valor

¹Como se verá más adelante, muestras con reemplazo de la muestra original.

particular y_i se calcula como:

$$P(Y^* = y_i) = \hat{F}_n(y_i) - \lim_{x \rightarrow y_i^-} \hat{F}_n(x).$$

Consideremos los siguientes dos casos:

- Caso $i = 1$. Como no hay valores menores que y_1 , la función de distribución empírica es cero justo antes de este punto:

$$\lim_{x \rightarrow y_1^-} \hat{F}_n(x) = 0.$$

Por lo tanto:

$$P(Y^* = y_1) = \hat{F}_n(y_1) = \frac{1}{n}.$$

- Caso $i > 1$: Justo antes de y_i , la distribución empírica ya ha acumulado la probabilidad de los valores anteriores:

$$\lim_{x \rightarrow y_i^-} \hat{F}_n(x) = \hat{F}_n(y_{i-1}).$$

Como cada observación recibe una probabilidad de $1/n$, tenemos:

$$\hat{F}_n(y_i) = \hat{F}_n(y_{i-1}) + \frac{1}{n}.$$

Por lo que:

$$P(Y^* = y_i) = \hat{F}_n(y_i) - \hat{F}_n(y_{i-1}) = \frac{1}{n}.$$

□

Dado que la función de distribución empírica $\hat{F}_n(y)$ asigna una probabilidad de $1/n$ a cada una de las observaciones y_1, y_2, \dots, y_n , generar una muestra aleatoria de tamaño n a partir de $\hat{F}_n(y)$ es equivalente a realizar el siguiente experimento con una urna:

1. Se colocan las observaciones y_1, y_2, \dots, y_n en la urna.
2. Se revuelven las observaciones.
3. Se extrae una observación al azar, registrando su valor.
4. Se devuelve la observación a la urna y se revuelve nuevamente el contenido.
5. Se repite el procedimiento hasta haber registrado n observaciones.

Dado que en cada extracción todas las observaciones tienen la misma probabilidad de ser seleccionadas ($1/n$) y que las extracciones son independientes unas de otras, este mecanismo de selección preserva las propiedades de una muestra aleatoria cuando la selección se realiza a partir de un número finito de elementos. En el contexto de poblaciones finitas, este procedimiento es conocido como muestreo aleatorio simple con reemplazo.

3. Estimación de la Varianza vía Bootstrap

Para comprender mejor el uso del Bootstrap, trabajaremos con uno de los ejemplos más simples, en el que todas las cantidades pueden calcularse analíticamente. Comenzamos, como siempre, asumiendo nuestro modelo de generación de datos:

Sea Y_1, \dots, Y_n una muestra aleatoria de $Y \sim F_Y(y)$, donde nos interesa estimar $\mu = \mathbb{E}(Y)$, suponiendo que $\sigma^2 = \text{Var}(Y) < \infty$. En este caso, el estimador plug-in es la media muestral:

$$\hat{\mu}_n = \bar{Y}_n.$$

Dado que estamos en uno de los casos más sencillos, es posible calcular analíticamente la varianza de la media muestral:

$$V(\hat{\mu}_n) = \frac{1}{n} V(Y) = \frac{\sigma^2}{n}, \quad (6)$$

donde observamos que la varianza del estimador depende directamente de la varianza de la variable aleatoria Y , la cual, a su vez, es un funcional estadístico (una función de F_Y). Para hacer esta relación explícita, escribimos:

$$V_{F_Y}(\hat{\mu}_n) = \frac{\sigma^2}{n}.$$

Sin embargo, en la práctica, si no conocemos μ , tampoco conocemos σ^2 , por lo que debemos estimarla mediante el estimador plug-in, es decir, usando la distribución empírica $\hat{F}_n(y)$. Así, obtenemos:

$$V_{\hat{F}_n}(\hat{\mu}_n) = \frac{\hat{\sigma}_n^2}{n}.$$

3.1. Estimación mediante Bootstrap

¿Cómo opera el Bootstrap en este caso? Procedemos generando múltiples muestras bootstrap de tamaño n , extraídas de $\hat{F}_n(y)$, y calculamos la media en cada una:

$$\begin{aligned} Y_1^{*1}, Y_2^{*1}, \dots, Y_n^{*1} \stackrel{\text{m.a.}}{\sim} \hat{F}_n(y), & \rightarrow \hat{\mu}_n^{*1} = \bar{Y}_n^{*1}, \\ Y_1^{*2}, Y_2^{*2}, \dots, Y_n^{*2} \stackrel{\text{m.a.}}{\sim} \hat{F}_n(y), & \rightarrow \hat{\mu}_n^{*2} = \bar{Y}_n^{*2}, \\ & \vdots \\ Y_1^{*B}, Y_2^{*B}, \dots, Y_n^{*B} \stackrel{\text{m.a.}}{\sim} \hat{F}_n(y), & \rightarrow \hat{\mu}_n^{*B} = \bar{Y}_n^{*B}. \end{aligned}$$

A partir de estas muestras bootstrap, la varianza de $\hat{\mu}_n$ se estima como:

$$V_{\text{Boot}}(\hat{\mu}_n) = \frac{1}{B} \sum_{j=1}^B (\hat{\mu}_n^{*j} - \bar{\hat{\mu}}_n^*)^2, \quad (7)$$

donde

$$\bar{\hat{\mu}}_n^* = \frac{1}{B} \sum_{j=1}^B \hat{\mu}_n^{*j}.$$

3.2. Comparación de Aproximaciones

Lo que obtenemos es la siguiente relación:

$$V_{F_Y}(\hat{\mu}_n) \underset{\text{no tan precisa}}{\approx} V_{\hat{F}_n}(\hat{\mu}_n) \underset{\text{más precisa}}{\approx} V_{Boot}(\hat{\mu}_n).$$

Nuestro objetivo es aproximar $V_{F_Y}(\hat{\mu}_n)$. Como se sugiere en la cadena de aproximaciones, la segunda aproximación mediante Bootstrap suele ser bastante precisa, especialmente si el número de re-muestras es grande ($B > 1,000$). Sin embargo, la precisión de la primera aproximación depende de qué tan bien \hat{F}_n aproxima a F_Y , lo cual está influenciado por el tamaño muestral n .

Además, la calidad de la estimación depende de la forma del estimador. En nuestro ejemplo, $\hat{\mu}_n$ es un estimador simple y estable, pero en problemas más generales, $\hat{\theta}_n$ podría ser una función más compleja de la muestra, lo que puede introducir inestabilidad numérica o sensibilidad a valores atípicos. En tales casos, la elección del método de remuestreo y el número de iteraciones juegan un papel clave en la confiabilidad del Bootstrap.

4. Intervalos de Confianza Bootstrap

Existen varias formas de construir intervalos de confianza usando Bootstrap. Aquí discutimos tres métodos.

4.1. Intervalo Normal

El método más sencillo es, primero asumir que

$$Z_n = \frac{\hat{\theta}_n - \theta}{se(\hat{\theta}_n)} \xrightarrow{D} Z \sim N(0, 1),$$

con $se(\hat{\theta}_n) = \sqrt{V(\hat{\theta}_n)}$, y entonces simplemente aproximar $V(\hat{\theta}_n)$ con $V_{Boot}(\hat{\theta}_n)$. Bajo estos supuestos el intervalo aproximado del $(1 - \alpha) \times 100\%$ de confianza estaría dado por:

$$(\hat{\theta}_n - z_{\alpha/2} se_{Boot}, \hat{\theta}_n + z_{\alpha/2} se_{Boot})$$

donde $se_{Boot} = \sqrt{V_{Boot}(\hat{\theta}_n)}$ es la estimación bootstrap del error estándar. Obviamente, este intervalo solo es preciso si la distribución de $\hat{\theta}_n$ es aproximadamente normal.

Método 2: Intervalos Pivotales

Sea $\theta = T(F)$ y $\hat{\theta}_n = T(\hat{F}_n)$, y definamos el pivote:

$$R_n = \hat{\theta}_n - \theta.$$

Sea $\hat{\theta}_n^{*1}, \dots, \hat{\theta}_n^{*B}$ el conjunto de réplicas bootstrap de $\hat{\theta}_n$. Definimos la función de distribución acumulada del pivote:

$$H(r) = P_F(R_n \leq r).$$

El intervalo pivotal se define como $C_n = (a, b)$, donde:

$$a = \hat{\theta}_n - H^{-1}(1 - \alpha/2), \quad b = \hat{\theta}_n - H^{-1}(\alpha/2).$$

Para demostrar que este intervalo satisface:

$$P_F(a \leq \theta \leq b) = 1 - \alpha,$$

primero sustituimos

$$P_F(\hat{\theta}_n - H^{-1}(1 - \alpha/2) \leq \theta \leq \hat{\theta}_n - H^{-1}(\alpha/2)).$$

y posteriormente restamos $\hat{\theta}_n$ en todas las partes de la desigualdad:

$$P_F(-H^{-1}(1 - \alpha/2) \leq \theta - \hat{\theta}_n \leq -H^{-1}(\alpha/2)).$$

Multiplicando por un menos, y considerando que $R_n = \hat{\theta}_n - \theta$, podemos reescribir:

$$\begin{aligned} P_F(H^{-1}(\alpha/2) \leq R_n \leq H^{-1}(1 - \alpha/2)) &= P_F(R_n \leq H^{-1}(1 - \alpha/2)) - P_F(R_n \leq H^{-1}(\alpha/2)) \\ &= H(H^{-1}(1 - \alpha/2)) - H(H^{-1}(\alpha/2)) \\ &= 1 - \alpha/2 - \alpha/2 \\ &= 1 - \alpha. \end{aligned}$$

Sin embargo, la función H es desconocida, por lo que estimamos su versión empírica mediante:

$$\hat{H}(r) = \frac{1}{B} \sum_{b=1}^B \mathbb{I}(R_n^* \leq r)$$

donde $R_n^* = \hat{\theta}_n^* - \hat{\theta}_n$: entonces estamos tratando de describir la variabilidad de $\hat{\theta}_n$ mediante las estimaciones Bootstrap $\hat{\theta}_n^*$ y a θ lo estimamos con $\hat{\theta}_n$. Definimos los cuantiles muestrales:

$$\hat{a} = \hat{\theta}_n - \hat{H}^{-1} \left(1 - \frac{\alpha}{2} \right) = \hat{\theta}_n - (\hat{\theta}_{n,1-\alpha/2}^* - \hat{\theta}_n),$$

$$\hat{b} = \hat{\theta}_n - \hat{H}^{-1} \left(\frac{\alpha}{2} \right) = \hat{\theta}_n - (\hat{\theta}_{n,\alpha/2}^* - \hat{\theta}_n).$$

Finalmente, el intervalo de confianza bootstrap pivotal es:

$$C_n = \left(2\hat{\theta}_n - \hat{\theta}_{n,1-\alpha/2}^*, 2\hat{\theta}_n - \hat{\theta}_{n,\alpha/2}^* \right).$$

en donde para calcular los cuantiles

1. Se ordenan los valores bootstrap de manera ascendente:

$$\hat{\theta}_n^{*(1)} \leq \hat{\theta}_n^{*(2)} \leq \dots \leq \hat{\theta}_n^{*(B)}.$$

2. Determina los cuantiles requeridos:

- El cuantil $\alpha/2$ se obtiene como:

$$\hat{\theta}_{n,\alpha/2}^* = \hat{\theta}_n^{*([\lfloor B \cdot \alpha/2 \rfloor])}$$

- El cuantil $1 - \alpha/2$ se obtiene como:

$$\hat{\theta}_{n,1-\alpha/2}^* = \hat{\theta}_n^{*([\lfloor B \cdot (1-\alpha/2) \rfloor])}$$

Método 3: Intervalos Percentiles

El intervalo percentil bootstrap se define como:

$$C_n = (\theta_{n,\alpha/2}^*, \theta_{n,1-\alpha/2}^*).$$

5. Ejercicios

5.1. Ejercicio 4

Sean X_1, \dots, X_n observaciones distintas. Demuestre que hay $\binom{2n-1}{n}$ muestras bootstrap distintas.

Solución

Cada muestra bootstrap puede representarse por un vector de conteos (N_1, \dots, N_n) , donde N_j indica cuántas veces se seleccionó X_j en la muestra bootstrap. Estos valores cumplen:

$$N_1 + N_2 + \cdots + N_n = n$$

El número de maneras de distribuir n elementos entre n categorías con repetición se obtiene mediante combinaciones con repetición:

$$\binom{n + (n - 1)}{n} = \binom{2n - 1}{n}$$

Por lo tanto, el número total de muestras bootstrap distintas es $\binom{2n-1}{n}$, completando la demostración.

Ejercicio 5

Sean X_1, \dots, X_n observaciones distintas (sin empates). Sea X_1^*, \dots, X_n^* una m.a. de $X^* \sim \hat{F}_n(x)$, define $\bar{X}_n^* = \frac{1}{n} \sum_{i=1}^n X_i^*$. Encuentra:

1. $E(\bar{X}_n^* | X_1, \dots, X_n)$
2. $V(\bar{X}_n^* | X_1, \dots, X_n)$
3. $E(\bar{X}_n^*)$
4. $V(\bar{X}_n^*)$

Solución

Cada X_i^* toma valores de X_j con probabilidad $1/n$, por lo que su esperanza condicional es:

$$E(X_i^* | X_1, \dots, X_n) = \sum_{j=1}^n X_j P(X_i^* = X_j) = \sum_{j=1}^n X_j \frac{1}{n} = \bar{X}_n$$

Dado que \bar{X}_n^* es la media de los X_i^* :

$$E(\bar{X}_n^* | X_1, \dots, X_n) = \bar{X}_n$$

Para la varianza condicional:

$$\begin{aligned} V(X_i^* | X_1, \dots, X_n) &= E(X_i^{*2} | X_1, \dots, X_n) - E(X_i^* | X_1, \dots, X_n)^2, \\ &= \sum_{j=1}^n X_j^2 P(X_i^* = X_j) - \bar{X}_n^2, \\ &= \frac{1}{n} \sum_{j=1}^n X_j^2 - \bar{X}_n^2 = S^2, \end{aligned}$$

donde

$$S^2 = \frac{1}{n} \sum_{j=1}^n (X_j - \bar{X})^2$$

Finalmente,

$$V(\bar{X}_n^* | X_1, \dots, X_n) = \frac{S^2}{n}$$

Por la esperanza iterada:

$$E(\bar{X}_n^*) = E(E(\bar{X}_n^* | X_1, \dots, X_n)) = E(\bar{X}) = \mu$$

Para la varianza total:

$$\begin{aligned} V(\bar{X}_n^*) &= E(V(\bar{X}_n^* | X_1, \dots, X_n)) + V(E(\bar{X}_n^* | X_1, \dots, X_n)) \\ &= E\left(\frac{S^2}{n}\right) + V(\bar{X}_n) \\ &= \frac{2\sigma^2}{n}. \end{aligned}$$

El último paso se debe a que $E(S^2) = \sigma^2$ y $V(\bar{X}_n) = \frac{\sigma^2}{n}$.

A. Bootstrap Bayesiano para Poblaciones Finitas

Partiendo de Efron (1979), Lo (1988) describe las ideas centrales del Bootstrap en poblaciones finitas en un contexto Bayesiano (que no hemos visto). Dado $y_{1:n}$, una muestra aleatoria simple sin reemplazo (SRSWOR) de una población finita $y_{1:N}$, nos interesa el parámetro poblacional finito:

$$\eta_N = \eta(y_{1:N}).$$

Así, la parte faltante de la población se imputa mediante:

$$\begin{aligned} Y_{n+1} &\sim \hat{F}_n(y), \\ Y_{n+2} &\sim \hat{F}_{n+1}(y) = \frac{1}{n+1} \left(n\hat{F}_n(y) + \mathbb{I}_{\{Y_{n+1} \leq y\}} \right), \\ Y_{n+3} &\sim \hat{F}_{n+2}(y) = \frac{1}{n+2} \left((n+1)\hat{F}_{n+1}(y) + \mathbb{I}_{\{Y_{n+2} \leq y\}} \right), \\ &\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ Y_N &\sim \hat{F}_{N-1}(y) = \frac{1}{N-1} \left((N-2)\hat{F}_{N-2}(y) + \mathbb{I}_{\{Y_{N-1} \leq y\}} \right). \end{aligned}$$

Una vez imputada la población finita, para realizar inferencia bayesiana calculamos:

$$\hat{\eta}_N = \eta(y_{1:n}, Y_{n+1:N}),$$

y la cuantificación de la incertidumbre se obtiene imputando B poblaciones completas, lo que genera $\hat{\eta}_N^1, \hat{\eta}_N^2, \dots, \hat{\eta}_N^B$. Los cuantiles correspondientes de esta muestra aleatoria se utilizan para cuantificar la incertidumbre en la estimación de η_N . Este método se denomina Bootstrap Bayesiano para Poblaciones Finitas (FPBB, por sus siglas en inglés).

El objetivo del FPBB es aproximar la distribución predictiva $p(\eta_N \mid y_{1:n})$, mientras que los frequentistas buscan obtener la distribución muestral de un estimador.

Siguiendo ideas frequentistas, primero necesitamos un estimador de η_N basado en la muestra $y_{1:n}$, denotado como $\hat{\eta}_n$. Una vez imputada la población finita $\{y_{1:n}, Y_{n+1:N}\}$, se selecciona una muestra aleatoria simple sin reemplazo $y_{1:n}^*$ y se calcula el estimador $\hat{\eta}_n^{*1}$. Este proceso se repite B veces para obtener $\hat{\eta}_n^{*1}, \dots, \hat{\eta}_n^{*B}$, y estos valores se utilizan para aproximar la distribución muestral de $\hat{\eta}_n$.

Estas son las ideas fundamentales detrás de los métodos bootstrap bayesianos y frequentistas para poblaciones finitas. Para concluir, es importante mencionar que Lo describió el FPBB en términos de una muestra de Pólya de tamaño $N - n$ tomada de la urna de datos $y_{1:n}$. Aunque nuestra descripción es diferente en términos de formulación, es esencialmente equivalente y lleva a una conclusión clave sobre la actualización de la función de distribución empírica.

El modelo de **urna de Pólya** describe un esquema de muestreo en el que, en lugar de extraer elementos de una población fija con reemplazo, cada vez que se extrae un elemento, este se devuelve junto con una copia adicional. De este modo, las observaciones más frecuentes tienen una mayor probabilidad de ser seleccionadas en el futuro, generando una estructura de dependencia entre las extracciones.

En el contexto del FPBB, este proceso es análogo a la forma en que la función de distribución empírica se actualiza iterativamente al imputar valores faltantes. En particular, la regla de actualización de $\hat{F}_n(y)$ es equivalente a un mecanismo donde la “urnas” de la muestra observada se va enriqueciendo con cada nueva imputación, lo que captura la incertidumbre sobre la población completa. Esta interpretación refuerza la conexión entre el FPBB y la inferencia bayesiana, al modelar la distribución de la población no observada de una manera adaptativa y coherente con la información disponible en la muestra.

Referencias

- Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife, *The Annals of Statistics* **7**(1): 1 – 26.
- Lo, A. Y. (1988). A Bayesian Bootstrap for a Finite Population, *The Annals of Statistics* **16**(4): 1684 – 1695.

Wasserman, L. (2004). *All of Statistics: A Concise Course in Statistical Inference*, Springer Texts in Statistics, Springer, New York.