Sobre el teorema de transformación integral de probabilidad

Jorge de la Vega

27 de agosto de 2018

1. Introducción

Esta nota revisa el teorema de transformación integral de probabilidad, que es utilizado para obtener muestras aleatorias de variables aleatorias generales a traves de variables aleatorias uniformes.

En el caso de variables aleatorias continuas con funciones de distribución estríctamente crecientes, el teorema es muy fácil de demostrar. Las complicaciones surgen cuando la función de distribución es creciente con intervalos de constancia (i.e. cuando no es estrictamente creciente), o bien cuando tiene discontinuidades, o ambas, como en el caso de la Figura 1.

En esta nota haremos una revisión de las diferentes maneras de demostrar el resultado, tratando de hacerlo lo más intuitivio posible pero siendo rigurosos en la demostración.

2. Definiciones

Para poder considerar el caso más general, necesitamos una definición más amplia de la inversa de una función. Si F es estrictamente creciente, la inversa está bien definida si se cumple la condición

$$F^{-1}(y) = x \iff F(x) = y.$$

Sin embargo, si F es constante en algún intervalo (como en el intervalo (x_2, x_3) en la Figura 1), entonces la condición no define adecuadamente a la inversa. Se requiere modificar la definición a la de una inversa generalizada.

Definición 1. *La inversa generalizada de F (o función cuantíl) se define como*

$$F^{-1}(y) = \text{ínf}\{x : F(x) \geq y\}.$$

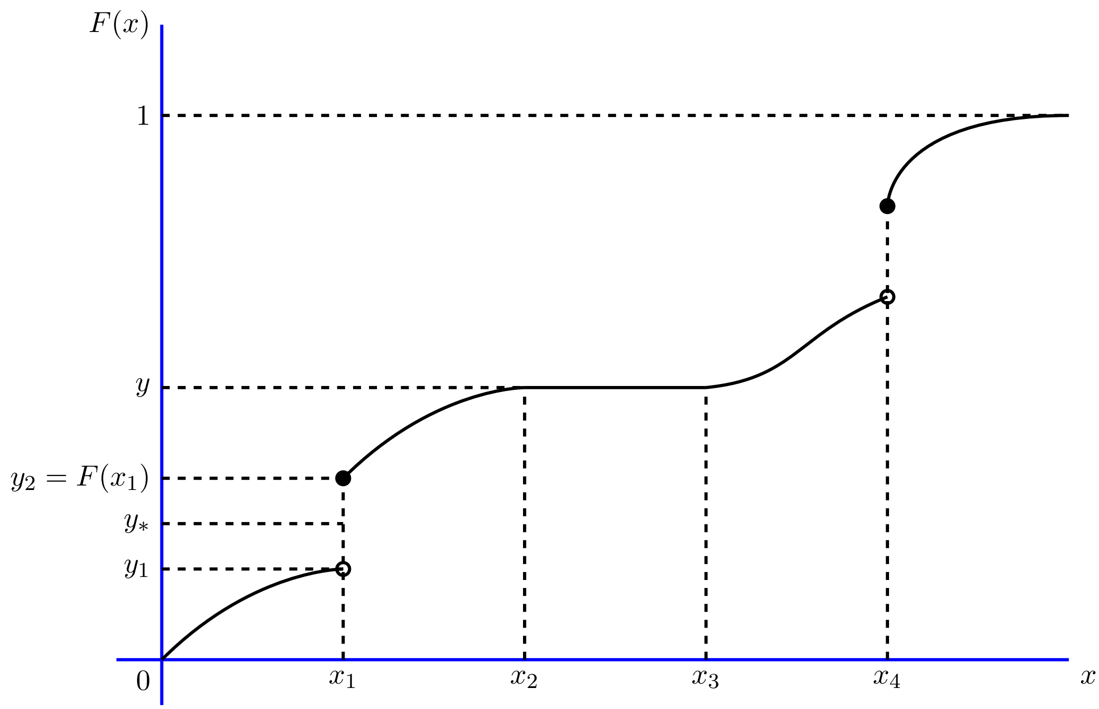

Figura 1: Función de distribución creciente y no continua

De acuerdo a esta definición, en la Figura 1, la observación y entonces está bien definida con esta inversa generalizada a ser $F^{-1}(y) = x_2$. En esa misma Figura para cualquier punto $y_* \in [y_1, y_2]$, se tiene $F^{-1}(y) = \text{ínf}\{x : F(x) \geq y\} = x_1$.
Para facilitar la comprensión, hagamos una gráfica de la función inversa generalizada de la Figura 1 para analizar su comportamiento. La Figura 2 muestra el resultado con los mismos puntos de referencia de la Figura 1.

De la gráfica de la función inversa, es importante notar las siguientes características y convencerse de ellas:

- La función cuantíl F^{-1} es no decreciente.
- Saltos en la función F se traducen a segmentos planos o de constancia en F^{-1}
- Regiones planas o de constancia en F se traducen a saltos en la función F^{-1}
- F es continua por la derecha, mientras que F^{-1} es continua por la izquierda.

3. Teorema de transformación integral de probabilidad

Aquí revisaremos la manera que se ha enunciado el teorema de transformación integral de probabilidad en diferentes textos.

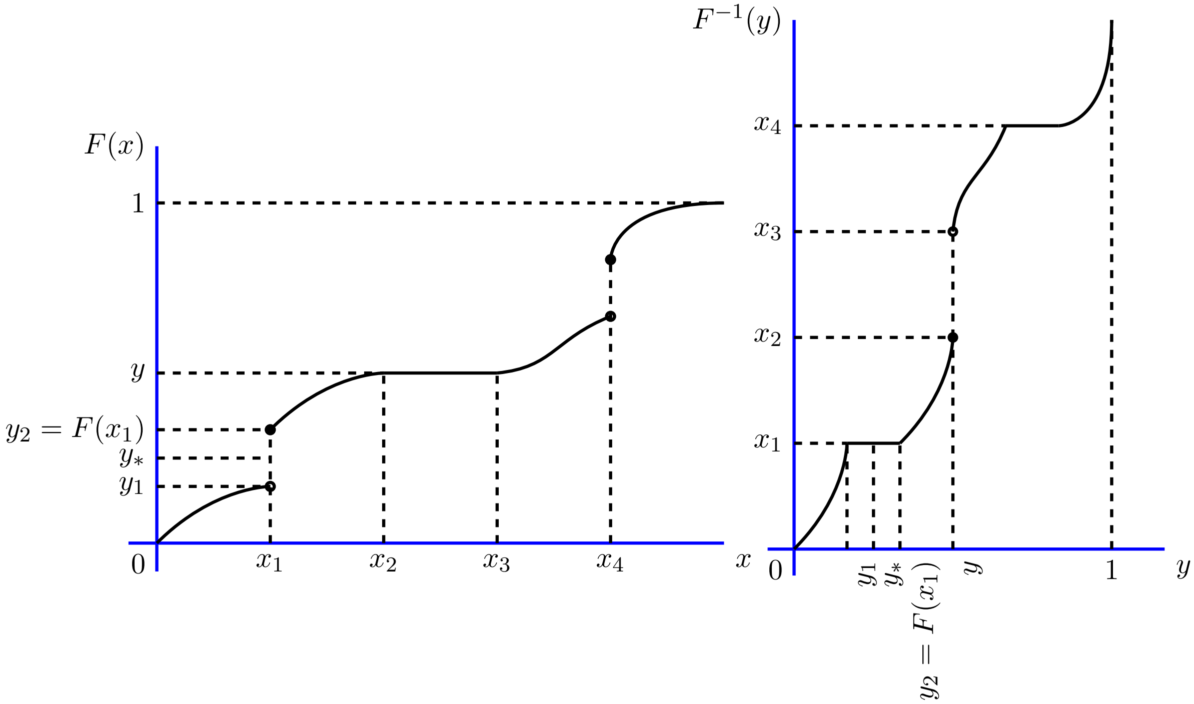

Figura 2: Función de distribución y su función cuantíl o generalizada inversa

En el libro de Casella y Berger [1], teorema 2.1.10, se muestra para el caso continuo.

Teorema 1. ***Teorema de transformación integral de probabilidad*** *Sea X una variable aleatoria que tiene función de distribución continua $F_X(x)$ y definimos la variable aleatoria Y como $Y = F_X(X)$. Entonces Y tiene distribución uniforme en el intervalo $(0, 1)$, $P(Y \leq y) = y$, $0 < y < 1$.*

Demostración. Para $Y = F_X(X)$ y $0 < y < 1$,

$$\begin{aligned} P(Y \leq y) &= P(F_X(X) \leq y) && (1) \\ &= P(F_X^{-1}(F_X(X)) \leq F_X^{-1}(y)) && (2) \\ &= P(X \leq F_X^{-1}(y)) && (3) \\ &= F_X(F_X^{-1}(y)) && (4) \\ &= y && (5) \end{aligned}$$

La igualdad (2) se cumple porque F_X^{-1} es creciente.

La igualdad (3), viene de la siguiente consideración: si F_X es estrictamente creciente, entonces claramanete se cumple que $F^{-1}(F_X(X)) = x$. Sin embargo, como ya se mencionó antes, en las partes planas de F_X, puede ser que $F_X^{-1}(F_X(x)) \neq x$. Supongamos que F_X es como en la Figura 1, y sea $x \in [x_2, x_3]$. Entonces $F_X^{-1}(F_X(x)) = x_2$ para cualquier x en ese intervalo. Pero aun en este caso, la probabilidad no se altera, porque $F_X(x) = F_X(x_1) = y$

para cualquier $x \in [x_2, x_3]$. La región plana denota un área donde la probabilidad es 0, ($P(x_2 < X \leq x) = F_X(x) - F_X(x_2) = 0$).

La igualdad (4) es la definición de F_X, y la (5) se da por la continuidad de F_X. $\square$

Una manera alternativa de probar el resultado propuesta por Angus [3] es la siguiente, pero se basa en un lema previo

Demostración. Consideremos el siguiente lema:

Lema. *Si X tiene función de distribución F, entonces para $x \in \mathbb{R}$,*

$$P(F(X) \leq F(x)) = F(x).$$

Demostración. El evento $\{F(X) \leq F(x)\}$ se puede descomponer de la siguiente manera:

$$\{F(X) \leq F(x)\} = [\{F(X) \leq F(x)\} \cap \{X \leq x\}] \cup [\{F(X) \leq F(x)\} \cap \{X > x\}].$$

Como $\{X \leq x\} \subset \{F(X) \leq F(x)\}$ y $\{X > x\} \cap \{F(X) < F(x)\} = \emptyset$, se sigue que:

$$\{F(X) \leq F(x)\} = \{X \leq x\} \cup [\{F(X) = F(x)\} \cap \{X > x\}].$$

Tomando probabilidades de estos eventos

$$P(F(X) \leq F(x)) = P(X \leq x) + P(F(X) = F(x), X > x) = P(X \leq x).$$

La última probabilidad es 0, porque X está en un intervalo donde F es constante. $\square$

Ahora se probará el Teorema 1. Sea $u \in (0, 1)$. Como F es continua, existe un número real x tal que $F(x) = u$. Entonces, por el lema anterior, $P(Y \leq u) = P(F(X) \leq F(x)) = u$. Esto implica que Y es uniforme. $\square$

La última igualdad (5) de la primera demostración anterior es la que tenemos que desarrollar para el caso de una variable X discreta, ya que su función de distribución tendrá discontinuidades. En el caso discreto, o mixto (con partes continuas y discretas), no se cumple que $Y = F_X(X)$ sea uniforme, pero se sigue cumpliendo que $F_X^{-1}(U) \sim F_X$, así que podemos seguir utilizando la transformación inversa sobre uniformes.

Para generalizar el teorema de la transformación inversa al caso más general posible, John E. Angus [3] propone el siguiente teorema, que se recomienda ir siguiendo en la Figura 2, y notando las características de la función cuantíl que ya se mencionaron previamente.

Teorema 2. ***Teorema de la función cuantil*** *Sea F una función de distribución X. Si se define F^{-1} como la función cuantíl $F^{-1}(y) = \text{ínf}\{x : F(x) \geq y\}, \quad 0 < y < 1$ y U tiene distribución uniforme en $(0, 1)$, entonces $X = F^{-1}(U)$ tiene función de distribución F.*

Demostración. Para cualquier x tal que $0 < F(x) < 1$ y cualquier $u \in (0,1)$ se cumple que

$$F(x) \geq u \iff x \geq F^{-1}(u)$$

Para ver esto, suponer que $x \geq F^{-1}(u) = \text{ínf}\{x : F(x) \geq u\}$. Como F es no decreciente y continua por la derecha, el conjunto $A_u = \{x : F(x) \geq u\}$ es un intervalo de la forma $[u, \infty)$. Por lo tanto, x debe satisfacer $F(x) \geq u$.

Conversamente, supongamos que $F(x) \geq u$. Entonces $x \geq \text{ínf}\{x : F(x) \geq u\} = F^{-1}(u)$. Se sigue entonces que

$$P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$$

$\square$

4. Otra demostración menos clara

Otra demostración, dirigida a una audiencia más avanzada, es la que se muestra en [2] (Capítulo 2, p. 38, proposición 2.2). Pero el resultado es el mismo. Se lista aquí para comparar con los argumentos dados antes. La demostración se muestra tal cual la redacta el autor.

Teorema 3. *Si F es la función de distribución de la variable aleatoria X, y $0 < u < 1$ entonces se cumple:*

(a) $u \leq F(x) \iff F^{-1}(u) \leq x$

(b) Si $U \sim \mathcal{U}(0,1)$, entonces $F^{-1}(U) \sim F$

(c) Si F es continua, entonces $F(X) \sim \mathcal{U}(0,1)$

Demostración. La parte (a) se sigue directamente de la definición de F^{-1} (es claro que $F(x) \geq u$ para $x \geq F^{-1}(u)$ pero no para cualquier $x < F^{-1}(u)$), y (b) sigue trivialmente de (a), lo que da:

$$P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x).$$

Finalmente para (c), tenemos que probar $P(F(X) \geq u) = 1 - u$, $0 < u < 1$. Sin embargo, por (a) y la continuidad de F,

$$P(F(X) \geq u) = P(X \geq F^{-1}(u)) = P(X > F^{-1}(u)) = 1 - F(F^{-1}(u)),$$

Asi que se debe probar que $F(F^{-1}(u)) = u$. Aquí $F(F^{-1}(u)9 \geq u$ es claro de (a) con $x = F^{-1}(u)$. Sea $y_n < F^{-1}(u)$ con y_n convergiendo por abajo a $F^{-1}(u)$. Entonces $F(y_n) < u$ por (a), y por la continuidad de F, $F(F^{-1}(u)) \leq u$. $\square$

Referencias

[1] Casella, George y Roger L. Berger *Statistical Inference, Second Ed.* Duxbury, USA, 2002.

[2] Asmussen, Søren. *Stochastic Simulation: Algorithms and Analysis*. Springer, 2007.

[3] Angus, John E. "The Probability Integral Transform and related results", *SIAM Review* **36** (1994):652-654.