PDF of the Random Variable when its Distribution Function Changes after the Change Points

C. D. Nanda Kumar and S. Srinivasan

Department of Mathematics, B.S.Abdur Rahman University
Vandalur, Chennai 600 048, India

Copyright © 2014 C. D. Nanda Kumar and S. Srinivasan. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

This paper derives the probability density function of the random variable X when the random variable X changes from one distribution to another after the change points. Illustrations are provided.

Mathematics Subject Classification: 60A05, 60A99, 60E05

Key words: random variable, probability density function, cumulative distribution function, change point or truncation point

1. Preliminaries

According to the probability theory the cumulative distribution function (CDF) describes the probability that a real valued random variable X with a given probability distribution will be found at a value less than or equal to x . In the case of a continuous distribution, it gives the area under the probability density function (PDF) from $-\infty$ to x .

Definition 1.1

The cumulative distribution function of a real-valued random variable X is the function given by

$$F_X(x) = P(X \leq x)$$

where the right-hand side represents the probability that the random variable X takes on a value less than or equal to x .

Definition 1.2

The probability that X lies in the interval (a, b) , where $a < b$, is therefore

$$P(a < X < b) = F_X(b) - F_X(a)$$

Definition 1.3

The CDF of the random variable X can be expressed as the integral of its PDF as follows.

$$F_X(x) = P(X \leq x) = \int_{-\infty}^x f_X(x) dx$$

Definition 1.4

The complementary CDF of the random variable X is defined and denoted as follows.

$$\overline{F}_X(x) = P(X > x) = 1 - F_X(x).$$

Properties of CDF 1.5

1. Every CDF F of the random variable X is monotone non-decreasing and right continuous.
2. $\lim_{x \rightarrow -\infty} F(x) = 0$.
3. $\lim_{x \rightarrow \infty} F(x) = 1$.

Definition 1.6

If A and B are any two events, then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Similarly, if A, B and C are any three events, then

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C).$$

More generally, for the events $A_1, A_2, A_3 \dots A_n$, we have by inclusion-exclusion principle:

$$\begin{aligned}
P\left(\bigcup_{i=1}^n A_i\right) &= \sum_{i=1}^n P(A_i) \\
&\quad - \sum_{i_1 < i_2} P(A_{i_1} \cap A_{i_2}) \\
&\quad + \sum_{i_1 < i_2 < i_3} P(A_{i_1} \cap A_{i_2} \cap A_{i_3}) - \dots \\
&\quad + (-1)^n \sum_{i_1 < i_2 < \dots < i_{n-1}} P(A_{i_1} \cap \dots \cap A_{i_{n-1}}) \\
&\quad + (-1)^{n+1} P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_n})
\end{aligned}$$

This result can be proved by mathematical induction.

Remark 1.7

If the events are independent, then

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)P(A_2) \dots P(A_n)$$

Remark 1.8

Obviously, we can find that from the definition 1.1, the CDF is very much connected with the event that the random variable $X < x$.

2. Main Result

A distribution having One Change Point

Suppose the random variable X follows the PDF $h_1(x)$ with the CDF $H_1(x)$ when $0 < x < \tau_1$ and it follows the PDF $h_2(x)$ with the CDF $H_2(x)$ when $\tau_1 < x < \infty$. (Since the change of distribution occurs at τ_1 , it is called as the change point)

Using Remarks 1.7, 1.8 and definition 1.6 for any two events, the CDF of X can be given as follows.

$$H(x) = \begin{cases} H_1(x) & \text{if } 0 < x < \tau_1 \\ H_1(\tau_1) + H_2(x - \tau_1) - H_1(\tau_1) \cdot H_2(x - \tau_1) & \text{if } \tau_1 < x < \infty \end{cases} \quad (2.1)$$

Simplifying (2.1) we get the following form

$$H(x) = \begin{cases} H_1(x) & \text{if } 0 < x < \tau_1 \\ H_1(\tau_1) + \overline{H_1}(\tau) H_2(x - \tau_1) & \text{if } \tau_1 < x < \infty \end{cases} \quad (2.2)$$

By differentiating (2.2), we get the PDF $h(x)$ of X . It is given by

$$h(x) = \begin{cases} h_1(x) & \text{if } 0 < x < \tau_1 \\ \overline{H_1}(\tau_1)h_2(x - \tau_1) & \text{if } \tau_1 < x < \infty \end{cases} \quad (2.3)$$

A distribution having Two Change Points

Suppose the random variable X follows the PDF $h_1(x)$ with the CDF $H_1(x)$ when $0 < x < \tau_1$, it follows the PDF $h_2(x)$ with the CDF $H_2(x)$ when $\tau_1 < x < \tau_2$ and it follows the PDF $h_3(x)$ with the CDF $H_3(x)$ when $\tau_2 < x < \infty$.

Using Remarks 1.7, 1.8 and definition 1.6 for any three events, the CDF of X can be given as follows.

$$H(x) = \begin{cases} H_1(x) & \text{if } 0 < x < \tau_1 \\ H_1(\tau_1) + H_2(x - \tau_1) - H_1(\tau_1) \cdot H_2(x - \tau_1) & \text{if } \tau_1 < x < \tau_2 \\ H_1(\tau_1) + H_2(\tau_2 - \tau_1) + H_3(x - \tau_2) - H_1(\tau_1) \cdot H_2(\tau_2 - \tau_1) - H_2(\tau_2 - \tau_1)H_3(x - \tau_2) - H_1(\tau_1)H_3(x - \tau_2) + H_1(\tau_1)H_2(\tau_2 - \tau_1)H_3(x - \tau_2) & \text{if } \tau_2 < x < \infty \end{cases} \quad (2.4)$$

Simplifying (2.4), we get the following form

$$H(x) = \begin{cases} H_1(x) & \text{if } 0 < x < \tau_1 \\ H_1(\tau_1) + \overline{H_1}(\tau)H_2(x - \tau_1) & \text{if } \tau_1 < x < \tau_2 \\ H_1(\tau_1) + \overline{H_1}(\tau)H_2(\tau_2 - \tau_1) + \overline{H_1}(\tau_1)\overline{H_2}(\tau_2 - \tau_1)H_3(x - \tau_2) & \text{if } \tau_2 < x < \infty \end{cases} \quad (2.5)$$

By differentiating (2.5), we get the PDF $h(x)$ of X .

$$h(x) = \begin{cases} h_1(x) & \text{if } 0 < x < \tau_1 \\ \overline{H_1}(\tau_1)h_2(x - \tau_1) & \text{if } \tau_1 < x < \tau_2 \\ \overline{H_1}(\tau_1)\overline{H_2}(\tau_2 - \tau_1)h_3(x - \tau_2) & \text{if } \tau_2 < x < \infty \end{cases} \quad (2.6)$$

We verify that (2.6) is a PDF.

Clearly $h(x) \geq 0$ by the assumptions of $H_1(x)$, $H_2(x)$ and $H_3(x)$.

$$\text{Also } \int_0^\infty h(x)dx = \int_0^{\tau_1} h(x)dx + \int_{\tau_1}^{\tau_2} h(x)dx + \int_{\tau_2}^\infty h(x)dx$$

$$= \int_0^{\tau_1} h_1(x)dx + \int_{\tau_1}^{\tau_2} \overline{H_1}(\tau_1)h_2(x - \tau_1)dx + \int_{\tau_2}^\infty \overline{H_1}(\tau_1)\overline{H_2}(\tau_2 - \tau_1)h_3(x - \tau_2)dx$$

$$\begin{aligned}
&= (H_1(\tau_1) - H_1(0)) + \overline{H}_1(\tau_1)(H_2(\tau_2 - \tau_1) - H_2(0)) \\
&\quad + \overline{H}_1(\tau_1)\overline{H}_2(\tau_2 - \tau_1)(H_3(\infty) - H_3(0)) \\
&= (H_1(\tau_1) - 0) + \overline{H}_1(\tau_1)(H_2(\tau_2 - \tau_1) - 0) + \overline{H}_1(\tau_1)\overline{H}_2(\tau_2 - \tau_1)(1 - 0) \\
&= H_1(\tau_1) + \overline{H}_1(\tau_1)H_2(\tau_2 - \tau_1) + \overline{H}_1(\tau_1)\overline{H}_2(\tau_2 - \tau_1) \\
&= H_1(\tau_1) + \overline{H}_1(\tau_1)(H_2(\tau_2 - \tau_1) + \overline{H}_2(\tau_2 - \tau_1)) \\
&= H_1(\tau_1) + \overline{H}_1(\tau_1) \\
&= 1
\end{aligned}$$

Therefore, $\int_0^\infty h(x)dx = 1$.

Thus, $h(x)$ is PDF.

A distribution having ‘n-1’ Change Points

Suppose the random variable X happens to change its distribution ‘n-1’ times, with change points $\tau_1, \tau_2, \dots, \tau_{n-1}$ then its PDF can be generalized as follows.

$$h(x) = \begin{cases} h_1(x) & \text{if } 0 < x < \tau_1 \\ \overline{H}_1(\tau_1)h_2(x - \tau_1) & \text{if } \tau_1 < x < \tau_2 \\ \overline{H}_1(\tau_1)\overline{H}_2(\tau_2 - \tau_1)h_3(x - \tau_2) & \text{if } \tau_2 < x < \tau_3 \\ \dots & \dots \\ \overline{H}_1(\tau_1)\overline{H}_2(\tau_2 - \tau_1) \dots h_n(x - \tau_{n-1}) & \text{if } \tau_{n-1} < x < \infty \end{cases}$$

3. Illustration

a) A distribution having one change point

3.1 Suppose the random variable

$X \sim \exp(\theta_1)$ before τ and $X \sim \exp(\theta_2)$ after τ then using 2.3 its pdf can be given by

$$h(x) = \begin{cases} \theta_1 e^{-\theta_1 x}, & 0 < x < \tau \\ e^{-(\theta_1 - \theta_2)\tau} \theta_2 e^{-\theta_2 x}, & x \geq \tau \end{cases}$$

This pdf has been used in [3].

3.2 Suppose the random variable

$X \sim \exp(\theta_1)$ before τ and $X \sim \text{erlang } 2(\theta_2)$ after τ then using 2.3

3.3

its pdf can be given by

$$h(x) = \begin{cases} \theta_1 e^{-\theta_1 x}, & 0 < x < \tau \\ \theta_2^2 (x - \tau) e^{-\theta_2(x-\tau)} e^{-\theta_1 \tau}, & x \geq \tau \end{cases}$$

This pdf has been used in [2] and [3].

3.4 Suppose the random variable

 $X \sim \exp(\theta)$ before τ and $X \sim \text{gamma } 2(k)$ after τ then using 2.3

its pdf can be given by

$$h(x) = \begin{cases} \theta e^{-\theta x}, & 0 < x < \tau \\ \frac{e^{-\theta \tau} 2^k (x - \tau)^{k-1} e^{-2(x-\tau)}}{\Gamma k}, & x \geq \tau \end{cases}$$

This pdf has been used in [1].

b) A distribution having two change points

3.5 Suppose the random variable

 $X \sim \exp(\theta_1)$ before τ_1 , $X \sim \exp(\theta_2)$ after τ_1 and before τ_2 and $X \sim \exp(\theta_3)$ after τ_2 then using 2.6

its pdf can be given by

$$h(x) = \begin{cases} \theta_1 e^{-\theta_1 x}, & 0 < x < \tau_1 \\ \theta_2 e^{-\theta_1 \tau_1} e^{-\theta_2(x-\tau_1)}, & \tau_1 \leq x < \tau_2 \\ \theta_3 e^{-\theta_1 \tau_1} e^{-\theta_2(\tau_2-\tau_1)} e^{-\theta_3(x-\tau_2)}, & x \geq \tau_2 \end{cases}$$

4. Future Work

The future work of this paper is deriving the pdf of the random variable x by treating the change point as the random variable.

References

- [1]. C.D. Nandakumar, S. Srinivasan, and P.S. Sehi Uduman, Optimal Reserve Inventory between two machines when the Repair Time has change of distribution after a change point, International Mathematical Forum, 7(54) (2012), 2659-2668.

- [2]. S. Srinivasan, A. Sulaiman and R. Sathiyamoorthi, Optimal Reserve Inventory between two machines under SCBZ property of inter-arrival times between breakdowns, International Journal of Physical Sciences – Ultra Science, 19(2M) (2007), 261-266.
- [3]. R.Suresh Kumar, Shock Model when the Threshold has a Change of Distribution after a Change Point, Journal of Indian Acad. Math, 28(1) (2007), 73-84.

Received: November 1, 2013