I'm not robot
reCAPTCHA
Continue

Probability distribution function properties pdf

The cumulative distribution function (cdf) of any variable \(X\) is the \(F_X \F_X\F_X function, both discrete and continuous random variables have cdfs, although we do not focus on them in the modules of discrete random variables and they are clearer to use for continuous random variables. As noted in the Discrete Probability Distributions module, the use of lowercase \ (x\) as an argument is arbitrary here : if we type \(F_X(t)\), this will be the same function determined by the random variable \(X\). But it helps to associate the corresponding lowercase letter with the random variable we are considering. Cdf is determined for all real values \(x\), sometimes implicitly, not explicitly. In the example in the previous section, we took into account the random variable \(T_i\), the total time spent on Facebook from individual \(i\) to 18 years of age. This is a measurement of time, in years, which should be between 0 and 18. So we know that \(T_i\) cdf must be zero for each value of \(t<0\). This means that if \ (t<0\), then \(F_{T_i}(t) = \Pr(T_i \leq t) = 0\). At the other extreme, we know that \(T_i\) should be less than or equal to 18. So if \ (t \ geq 18 \), then \(F_{T_i}(t) = \Pr(T_i \leq t) = 1 \). Remember that random variable results determine events in the event space, which is why we are able to attach probabilities to such results. Let \(X\) be any variable with cdf \(F_X(x)\). For \ (a<b\) we can consider the following events: \(C =\) '\(X \leq a\)" \(D =\) \(D=\) \(< X \leq b\)" \(E =\) \'\(X\leq b\). Then \(C\) and \(D\) are mutually exclusive and their union is the \(E\) event. On a third axiom of probabilities, this tells us that \ begin { alignat *} {2} \Pr(E) & = \ Pr (C) + \ Pr (D) \ \ \ \\ > Quad > \Pr(X \leq b) > \Pr (X \leq a) + \ Pr (< X \leq b) \ \ *\\ \\ \\ \\ "" "" "< X \leq b) &= < F_X(b) - F_X(a). \end{ alignat*} The cumulative allocation function \(F_X(x)\) of a random variable \(X\) has three important properties: the cumulative distribution function \(F_X(x)\) is a non-decreasing function. This follows directly from the result you just received: For \(< b\), we have \ [o, < < X \leq b) \geq \\ \, F_X(b) - F_X(a) 0 \ \, \, F_X(a) F_X(b). \] As \(x \to -\infty\), the value of \(F_X(x)\) approaches 0 (or equals 0). This is partly because \(\Lim\limits_{ x\to-infty} F_X(x) = 0\). This is partly because \(\Pr(\Varnothing) = 0\). \(x\to\infty\), the value of \(F_X(x)\) is closer to 1 (or equals 1). \(\lim\limits_{ x\to\infty} F_X(x) = 1 \). This is partly due to the fact that \(\Pr(\mathcal{E}) = 1 \). All these discussions apply equally to discrete and continuous random variables. Its form something similar, which is shown in the following figure. We require random variable to have a cdf that is a continuous function. Detailed description: General type of cumulative distribution function of a continuous random variable. We are now using a CDF continuous random variable to start thinking about the question of probabilities for continuous random variables. For discrete random variables, probabilities come directly from the probability function \(p_X(x)\): we identify the possible discrete values that any variable may take, and then indicate in some way the probability of each of these values. For a continuous random variable \(X\), once we know that cdf \(F_X(x)\) we can detect the probability that \((X\) lies in each given interval: \[Pr (< X \leq b) = F_X(b) - F_X(). \] But if we are interested in the probability that a random variable is using one specific value? What is \(\Pr(X=x)\) for a continuous random variable? We can deal with this issue , taking into account the probability that \(X\) is in the interval, and then reducing the interval to one point. Formally for continuous random variable\(X\) with cdf \(F_X(x)\): \ begin{ alignat *} {2} \Pr (X = x) &lim_{lim_{to 0^} \Pr(x-h < X \leq x) \\ \\\lt; \\\\\\\\\lim_ \< \\\\\\\< \\F_X(x) - F_X(x)\big) \\\\\\\\F_X\\\\\x) - F_X(x) > &\qquad &{ because} F_X \text{ is continuous}) \\\\\\\\\\\0. \end{alignat*} Therefore, \(\Pr(X= x) = 0\). It's a somewhat disturbing result: It seems we can never observe a continuous random variable that accepts a certain value, since the probability of observing each value is zero. Actually, that's true. Continuous random variables are really abstractions: what we observe is always rounded in some way. Helps to think about some specific cases. Supposing we consider real numbers randomly selected between 0 and 1 for which we save \(X_1\), the random number cut to one decimal place. For example, the number 0.07491234008 is saved as 0.0 (the first decimal place is zero). Keep in mind that this means that we are not rounded, but a rupture. If the mechanism generating these numbers does not have a preference for any position in the interval \(0,1)\), then the distribution of numbers, that we receive will be such that \ [\Pr(X_1 = 0.0) = \ Pr (X_1 = 0.1) = \s = \pr (X_1 = 0.9) = \tfrac{1}{10}. \] This is a discrete random variable, with the same probability for each of the ten possible results. Now let's assume that instead we save \(X_2\), the random number is cut to two decimal numbers (again, not rounding). For example, if the real number is \(0.9790295134\), we save 0.97. This random variable is also discreet, with \[\Pr(X_2= 0.00) = \Pr(X_2 = 0.01) = \dots = \Pr(X_2= 0. 99) = \tfrac{1}{100}. \] Distributions of \(X_1\) and \(X_2\) are shown in Figure 2. Detailed description Figure Probability functions \(p_{X_1}(x)\) for \(X_1\) and \(p_{X_2}(x)\) for \(X_2\). You can see where it goes: If we record the first \(k\) decimal places, positions, \(X_k\) has \(10^k\) possible results, each with the same probability \(10^-k}\). Excel has a function that creates real numbers between 0 and 1 selected so that there is no preference for each position in the interval \(0,1)\). If you type \((sf\text{=RAND()}\) into a cell and press return, you will get this number. Increase the number of decimal places shown in the cell. Keep going until you get many zeros at the end of the number; you need to increase the cell size. Your spreadsheet should look like Figure 3, although of course the specific number will be different... it's accidental, after all! Figure 3: Excel spreadsheet with any number from 0 to 1. If you press the 'F9' key repeatedly at this point, you will see a series of random numbers, all between 0 and 1. From their examination, it appears that Excel actually produces observations on the random variable \(X_{15}\), the first 15 decimal places of the number. So the probability of a specific of these numbers is \(10^{-15}\). Now consider the probability that each of these discrete random variables \(X_1, X_2, X_3, \dots \) will take a value in the interval \(0.3,0.4)\), for example. We have \begin{align*} \Pr(0.3 \leq X_1 < 0.4) &= \Pr(X_1 = 0.3) = 0.1, \\\ \Pr(0.3 \leq X_2 < 0.4) &= \Pr(X_2 = 0.30) + \Pr(X_2 = 0.31) + \dots + \Pr(X_2 = 0.39) \\\ &= 10 \times 0.01 = 0.1, \\\ \Pr(0.3 \leq X_3 < 0.4) &= 10^2 \times 10^{-3} = 0.1, \\\ \Pr(0.3 \leq X_4 < 0.4) &= 10^3 \times 10^{-4} = 0.1 , \\\ &\vdots \\\ \Pr(0.3 \leq X_k < 0.4) &= 10^{k-1} \times 10^{-k} = 0.1, \\\ & \vdots \end{align*} As we make the distribution finer and finer, with more and more possible discrete values, the probability that any of these discrete random variables lies in the interval \([0.3,0.4)\) is always 0.1. As \((k\) increases, the discrete random variable\(X_k\) can be considered closer and closer to a continuous random variable. This continuous random variable, which we label \(U\), has the property\ [\Pr(a \leq U \leq b) = b - a, \qquad\ \text{for } 0 \leq a \leq b \leq 1. \] For example\ (\ Pr (0. 3 \leq U \leq 0,4) = 0, 4 - 0 , 3 = 0,3 = 0,3 \ Cdf of this random variable, \(F_U(u)\), is very simple. This is \[F_U(u) = \Pr(U \leq u) = \begin{cases} 0 &\ text{ if } u \leq 0, \\ \ \ u > text{ if } < u \leq 1, \\\ 1 &\ text if > 1. \end{cases} The random variable with this CDF is said to have the same interval allocation \(0,1)\); we mean this by \(U stackrel{\mathrm{d}}=} \mathrm{U}(0,1)\). The following figure shows the graph of the cumulative distribution function of \(U\). Figure 4: The cumulative distribution function of \(U\stackrel{\mathrm{d}}}}} \mathrm{U}(0,1)\). You may have noticed a difference in the way the interval between 0.3 and 0.4 was treated in a continuous case compared to discreet cases. In all individual cases, the \(0.4, 0.40, 0.400, \dots\) are excluded, while for the variable that is included. Every time it comes to individual random variables, whether inequality is severe or not is often important and care needs to be taken care of. For example\ [\Pr (0.30 \leq X_2 \leq 0.40) = 11 \times 0,01 = 0,11 eq 0.1. \] On the other hand, as we noted above in this section, for each continuous random variable \(X\), we have \(Pr(x = x) = 0\). Therefore, for a continuous random variable \(X\) and for \(\leq b\), we have \(\Pr(X=a) = \Pr(X=b) = 0\) and therefore \[\left.\ begin{array}{ l} (< X < b) \\\\\\\\\\\\\\\\\\\\\\\\\\\\F_X F_X\\\\\\\\\\\\\\\\\\\