I'm not robot

reCAPTCHA

Continue

# Cdf to pdf probability

Video Available 3.2.1 Cumulative PMF distribution function is one way to describe the distribution of a discrete random variable. As we will see later, PMF cannot be defined for continuous random variables. The random variable cumulative distribution function (CDF) is another method to describe the distribution of random variables. The advantage of CDF is that it can be defined for any kind of random variable (discrete, continuous and mixed). Definition The cumulative distribution function (CDF) of the random variable $X$ is defined as $$F_X(x) = P(X \leq x), \textrm{ for all }x \in \mathbb{R}.$$ Note that the subscript of $X$ indicates that it is a random $X$ CDF. Note also that CDF is defined for all $x \in \mathbb{R}$. Let's look at an example. An example I've twice tossed a coin. Let $X be the number of heads observed. Find the CDF $X$. Workaround Note that $X \sim Binomial (2, \frac{1}{2})$. The $X$$R_X=\{0,1,2\}$ and its PMF is given $$P_X(0)=P(X=0)=\frac{1}{4},$$ $$P _X(1) =P(X=1)=\frac{1}{2},$$$P_X(2)=P(X=2)=\frac{1}{4}.$$ To find the CDF, we argue as follows. First, note that if $x < 0$, then $$F_X(x)=P(X \leq x)=0, \textrm{ for } x < 0.$$ Next, if $x\geq 2$, $$F_X(x)=P(X \leq x)=1, \textrm{ for } x\geq 2.$$ Next, if $0 \leq x < 1$, $$F_X(x)=P(X \leq x)=P(X=0)=\frac{1}{4}, \textrm{ for } 0 \leq x < 1.$$ Finally if $1 \leq x < 2$, $$F_X(x)=P(X \leq x)=P(X=0)+P(X=1)=\frac{1}{4}+\frac{1}{2}=\frac{3}{4}, \textrm{ for } 1 \leq x < 2.$$ Thus, to summarize, we have \begin{equation} onumber F_X(x) = \left\{ \begin{array}{l} 0 & \quad \text{for } x < 0\\ \frac{1}{4} & \quad \text{for } 0 \leq x < 1\\ \frac{3}{4} & \quad \text{for } 1 \leq x < 2 \\ 1 & \quad \text{for } x \geq 2\\ \end{array} \right. \end{equation} Note that when you are asked to find the CDF random variable, you must find the function for the entire actual line. Also, for discrete random variables, we need to be careful when you use $<$or $\leq$. Figure 3.3 shows the $F_X(x)$. Note that the CDF is flat between the points in $R_X$ and jumps to each value in the range. The size of the jump at each point is equal to the probability at that point. For example, at $x =1$, cdf jumps from $\frac{1}{4}$ to $\frac{3}{4}$. The jump size here is $\frac{3}{4}-\frac{1}{4}=\frac{1}{2}$, which equals $P_X(1)$. Note also that the open and closed circles at $x=1$ indicate that $F_X(1)=\frac{3}{4}$ and not $\frac{1}{4}$. For example, the CDF 3.9. In general, let $$X be a discreet random variable with a range of $R_X =\{x_1,x_2,x_3,...\}$, so that $x_1 < x_2 < x_3 < ...$ Here, for simplicity, we assume that the range $R_X$ is bounded from below, i.e. $$x_1 is the smallest value in $R_X$. If not, then $F_X(x)$ approaches zero as $x \rightarrow -\infty$ rather than hitting zero. Figure 3.4 shows the general form of the CDF, $F_X(x)$, for such a random variable. We see that the CDF in the form of a staircase. Note in particular that the CDF starts at $0$; i.e., $F_X(-\infty)=0$. Then it jumps at every point in the range. In particular, the CDF will remain flat between $x_k$ and $x_{k + 1}$, so we can type $$F_X(x)=F_X(x_k), \textrm{ for }x_k \leq x < x_{k+1}.$$ CDF jumps on each $x_k$. Specifically, we can type $$F_X(x_k)-F_X(x_k-\epsilon)=P_X(x_k), \textrm{ For $\epsilon>0$ small enough.} $$ that CDF is always a non-decreasing function, i.e. if $y \geq x $ then $F_X(y)\geq F_X(x)$. Eventually, CDF approaches $1$ as $x$ becomes big. We can type $$\lim_{x \rightarrow \infty} F_X(x)=1.$$ Fig.3.4 - CDF discrete random variable. Note that CDF completely describes the distribution of discrete random variables. In particular, we can find PMF values by looking at the jump values in the CDF function. Also, if we have PMF, we can find the CDF from it. Especially if $R_X=\{x_1,x_2,x_3,...\}$, we can type $$F_X(x)=\sum_{x_k \leq x} P_X(x_k).$$ Now let's show a useful formula. For all $a \leq b$, we have $$\hspace{50pt} P(a < X \leq b)=F_X(b)-F_X(a) \hspace{80pt} (3.1)$$ to view this, Note that for $a \leq b$ we have $$P(X \leq b)=P(X \leq a) + P(a < X \leq b).$$ So, $$F_X(b)=F_X(a) + P(a < X \leq b).$$ Pay attention again to use $< $ and $\leq$ because they could change for discrete random variables. Later, we'll see that equation 3.1 applies to all types of random variables (discrete, continuous, and mixed). Note that CDF gives us a $P (X \leq x)$. To $P (X < x)$, for a discrete random variable, we can easily type $$P(X < x)=P(X \leq x)-P(X=x)=F_X(x)-P_X(x).$$Example Flight $X$ to be a discrete random variable with a range of $R_X={1,2,3,...\}$. Supposing $$X$ PMF is given $$P_X(k)=\frac{1}{2^k} \textrm{ for } k=1,2,3,...$$ Locate and eject the CDF of $$X,$F_X(x)$. Find $P(2 < X \leq 5)$. $P (X > 4)$. First, note that this is a valid PMF. In particular, $$\sum_{k=1}^{\infty} P_X(k)=\sum_{k=1}^{infty} \frac{1}{2^k}=1 \textrm{ (geometric sum)}$$ To search for a CDF, Note that $\textrm{Pro } x < 1.$ $F_X(x)=0$. $\textrm{Pro } 1\leq x < 2,$$F_X(x)=P_X(1)=\frac{1}{2}$. $\textrm{For } 2\leq x < 3,$ $F_X(x)=P_X(1)+P_X(2)=\frac{1}{2}+ \frac{1}{4}=\frac{3}{4}$. In general, we have $$\textrm{Pro } 0 < k \leq x < k+1,$$$$F_X(x) =P_X(1)+P_X(2)+...+P_X(k)$$$=\frac{1}{2}+ \fra c{1}{4}+...+\frac{1}{2^k}=\frac{2^k-1}{2^k}.$$ Figure 3.5 shows the CDF $X$. the CDF of the random variable given in example 3.10. To find $P(2 < X \leq 5)$, we can type $$P(2 < X \leq 5)=F_X(5)-F_X(2)=\frac{31}{32}-\frac{3}{4}=\frac{7}{32}.$$ Or equivalent, we can type $$P(2 < X \leq 5)=P_X(3)+P_X(4)+P_X(5)=\frac{1}{8}+\frac{1}{16}+\frac{1}{32}=\frac{7}{32} ,$$, which provides the same answer. To find $P(X > 4)$, we can write $$P(X > 4)=1-P(X \leq 4)=1-F_X(4)=1-\frac{15}{16}=\frac{1}{16}.$$ Recall that continuous random have countless possible values (think of real numbers). As with discrete random variables, we can talk about the probability of continuous random variables using density functions. The probability density (pdf) function, marked \(f\), a continuous random variable \(X\) meets the following: \(f(x) \geq 0\), for all \(x\in\mathbb{R}\) \(f\) is a review context \(\displaystyle{\int\limits^{\infty}_{-\infty}\! f(x)\,dx =1}\) \(\displaystyle{P(a\leq X\leq b) = \int\limits^a_b\! f(x)\,dx}\) The first three conditions in the definition list the properties needed for a function to be a valid pdf for a continuous random variable. The fourth condition tells us how to use pdf to calculate probability for continuous random variables that are given by integrals continuous analog to amounts. Let the random variable \(X\) indicate the time when the person is waiting for the elevator to arrive. Supposing the longest will have to wait for the elevator is 2 minutes, so the possible values \(X\) (in minutes) are given by the interval \([0,2]\). The possible pdf file for \(X\) is given $$f(x) = \left\{\begin{array}{l} x, & \text{for}\ 0\leq x\leq 1 \\ 2-x, & \text{for}\ 1< x\leq 2 \\ 0, & \text{otherwise} \end{array}\right.otag$$ Chart \(f\) is listed below to verify that \(f\) meets the first three conditions in definition 4.1.1: It is clear from the chart that \(f(x) \geq 0\) , for all \(x \in \mathbb{R}\). Since there are no holes, jumps, asymptotes, we see that \(f(x)\) is (piece) continuous. Finally, we calculate: $$\int\limits^{\infty}_{-\infty}\! f(x)\,dx =\int\limits^{2}_0\! x\,dx =\int\limits^1_0\! x\,dx +\int\limits^2_0\! (2-x)\,dx = 1otag$$ Figure 1: Pdf chart for \(X\), \(f(x)\) So, To calculate the probability that a person is waiting less than 30 seconds (or 0.5 minutes) for the arrival of the elevator, we calculate the following probability using pdf and fourth property in definition 4.1.1: $$P(0\leq X\leq 0.5) = \int\limits^{0.5}_0\! f(x)\,dx =\int\limits^{0.5}_0\! x\,dx = 0.125otag$$ Note that unlike discrete random variables, continuous random variables have a zero probability, i.e. the probability that a continuous random variable equals one value is always given by 0. Formally, this results from the properties of the integrals: $$P(X=a) = P(a\leq X\leq a) = \int\limits^a_a\! f(x)\, dx =0.otag$$ Informally, if we consider that the probability for a continuous random variable is given by the area below pdf, then because there is no area in the row, there is no probability associated with a random variable s per value. This does not mean that a continuous random variable is never equal to a single value, only that we do not assign any probability to individual random variable values. For this reason, we are only talking about the probability of a continuous random variable having a value in INTERVAL, not a point. And whether or not interval endpoints are included does not affect probability. In fact, pravdepodobnosť je rovnaká: $$P(a\leq X\leq b) = P(a<X<b) = P(a\leq X< b) = P(a< X \leq b) = \int\limits^b_a\!f(x)\,dxotag$$ \int\limits^b_a\!f(x)\,dxotag$$