

Say random variable with some range 0-R and say R is a whole number. Approximate it with buckets, discrete probability 1 Bucket, lone rectangle base R height 1, Area R Then R buckets where the i th bucket has height that is it's probability, $P([i, i + 1])$. Then the area is 1 since the heights of the bars are a probability distribution and thus sum to 1 and the widths of the bars are 1 because the a range of R spanned by R buckets means each bucket is 1 unit wide. Thus the area becomes $\sum_i 1 * P([i, i + 1]) = 1$.

From a scheme of 1 bucket to subdividing that lone bucket into R buckets, I want to point out two things. First, the histogram with R times more buckets generally gives a better sense of the 'shape' of the true distribution of the random variable. Second, the area of this histogram decreases by a factor of $\frac{1}{R}$. The first point seems obvious and suggests the strategy of taking more buckets of diminishing widths for finer resolution. And let me explore the second point, the differences of the area tied to the bucketing strategy.

So have some interval, $[x, x + \Delta x]$, of size Δx with a probability P . In a discrete probability distribution that is a histogram, this interval has an associated bar with area as base times height or $(\Delta x)P$. The base is the interval size or width Δx , and the height is the raw probability of the interval, $P([x, x + \Delta x]) = P$. I want to split this interval into 3 intervals, so $[x, x + \frac{\Delta x}{3}]$, $[x + \frac{\Delta x}{3}, x + 2 * \frac{\Delta x}{3}]$, $[x + 2 * \frac{\Delta x}{3}, x + \Delta x]$. The intervals' associated probabilities must sum to P then, the probability of the combined interval. And 3 intervals have 3 associated bars on the new, higher resolution, histogram. And I aim to show that the total area of these 3 bars is exactly $\frac{1}{3}$ rd the area of the original bar.

There are two ways I can do this, width and a height oriented ways, that both rely on the facts that the three bars have equivalent widths of $\frac{\Delta x}{3}$ and heights that sum to P .

First the width way. If you stack these three bars on top of each other, they will reach the height of the original bar, P . But the width of this construction is $\frac{\Delta x}{3}$, or $\frac{1}{3}$ rd the width of the original bar. TODO picture of this I can show this algebraically easily:

$$A = w * h_1 + w * h_2 + w * h_3 = w * \left(\sum_i h_i \right) \quad (1)$$

$$= \left(\frac{\Delta x}{3} \right) * P = \frac{\Delta x * P}{3} \quad (2)$$

Next the height way that is a different interpretation of the same algebra. Let me explore two different sets of size 3 probabilities these intervals might happen to take on. First, let me treat the random variable is uniform throughout $[x, x + \Delta x]$. Then these 3 subintervals would have probabilities $[\frac{P}{3}, \frac{P}{3}, \frac{P}{3}]$. Now if you view these bars as they are in the new histogram, that is side by side, you see that they form a rectangle with height $\frac{P}{3}$ and the same combined width as the original, Δx . TODO picture of this And thus the area is the desired $\frac{\Delta x * P}{3}$, this time the same width but $\frac{1}{3}$ rd height of the original bar. However note that there is no stipulation on uniformity. So the second set of 3 probabilities the 3 subintervals could take on, as I am inspired by the arithmetic sequence $P - (\Delta P), P, P + (\Delta P)$ that averages out to P , are probabilities $[\frac{P - (\Delta P)}{3}, \frac{P}{3}, \frac{P + (\Delta P)}{3}]$ for some small ΔP such that the first probability does not dip below 0. Verify these probabilities sum to the desired P . So I have 3 probabilitites with a fixed sum, P , so the average probability is then $\frac{P}{3}$. So if I view these new bars on the histogram side by side, their average height is $\frac{P}{3}$ though only the middle of the bars has that average height. TODO picture of this. And as the width of each of the bars is fixed, the total area is directly proportional to the total sum of the heights (with the width as the proportionality constant, see Equation 1), which when applying the definition of average to the bar

heights, gives me the statement that the total sum of the 3 bar heights is the average height multiplied by 3, the number of heights. So these 3 bars with an average height of $\frac{P}{3}$, stacked side by side as they are on the new histogram, are equivalent in total area as the uniform case where all the bars have the same uniform height that is that same average height of $\frac{P}{3}$. TODO Say previous 2 pictures are equal in total area of the bars

If this application of the definition of average is unclear, see my average notes work example where [1,1,3] division of work among 3 students could be more fairly [4/3, 4/3, 4/3] (TODO actually add this to said note).

The upshot is that now I am convinced if I have some bucketing scheme with N uniformly spread buckets, and then I subdivide each of the buckets into K more uniformly spaced buckets, the area of the histogram would get cut by a factor of $\frac{1}{K}$ as each bar would become split into K bars with a total of $\frac{1}{K}$ th the area of the old bar. And doing this for all N bars would result in a new area $\frac{1}{K}$ times the original one but with K times more buckets. So there's a tradeoff. The more buckets I have, I will get a more finer grained sense of the shape of the distribution, but the area would get cut into $1/K$. N to $N * K$ buckets, $\frac{R}{N}$ to $\frac{R}{KN}$ bucket width, A to $\frac{A}{K}$ area.

A resolution then that I can take is whenever I subdivide an interval into K intervals of $\frac{1}{K}$ the original interval width each, then I can take the height of each of the $\frac{1}{K}$ width intervals and multiply them by K . This way the area of the histograms would remain unchanged while the resolution would improve. So in the previous example, where $K = 3$ subdivisions, the average height or average probability from 0.1 gets scaled by K to 0.3, which is the original probability of the full $[x, x + \Delta x]$ interval. So the interval with some height P gets split into K intervals with $\frac{1}{K}$ the width but average height P meaning the sum of the areas of these K bars is unchanged. TODO it's late let me refine thoughts think have some "spark" Now let me say the height of each bar is not the probability of the backing interval of the bar but instead the height is directly proportional to the probability. So dividing that interval into K subintervals... fair to say the average heights of K bars would be the original height divided by K and thus proportional to the original interval's probability divided by K with same proportionality constant? Then multiplying all subintervals' heights by K would make their average height proportional to the original interval's probability with the same proportionality constant? And multiplying by K is the same as dividing by $\frac{1}{K}$ which is the width (in terms of intervals original width???) nope no sparks...something like elizabeth and step size with standard as 1 so R buckets...

num boxes = max weight capacity / box weight $n = 24/b$ where n is num boxes and b is box weight if $b=6$ then $n=4$ So can fit 4 boxes of weight 6 each to exactly fit in capacity 24. what if some technology induces b to decrease k -fold, say $k=3$ and b becomes b/k , so from box weight 6 to weight 2 Then you could fit 3 times as many boxes as per box previous fit, 3 of these new boxes could replace it $n = 24/(6/3) = 24/2=12$ and indeed 12 is 3 times 4, the number of boxes you used to be able to fit More generally $n = 24/(b/k) = k(24/b)$ **compensation, $b \rightarrow b/3$ then $n \rightarrow n*3$ can rearrange $n = 24/b$ to $n*b = 24$ so $(3*n)(b/3) = 24$ 24 is constant, constant capacity. Likewise want area in graph to be constant. $h*w = A$ where A is fixed let w be Width of bar. $w \rightarrow w/3$ then want $h \rightarrow h*3$**

Right now don't have a graph but have a histogram with a given step size. Even if slash width by 3, still a histogram. Want it to converge to a graph. Thought experiment. Say contour, skyline (leet code problem lol), forms a graph That graph alone makes no sense unless I specify an interval Say graph is uniform on $[x, x+\delta]$ Prob of that interval is height Or prob if $[x, x+\delta/3]$ is that height as well?

OK So interval with width Δx and some prob P . Split into K intervals each with width $\frac{\Delta x}{K}$. Then each of the intervals on average will have prob $\frac{P}{K}$ and thus the area concatenating(combining) these

bars together for span of the original Δx interval results in the total area going from $\Delta x * P$ to $\frac{\Delta x * P}{K}$, so to compensate and preserve the area as a fixed constant over this interval regardless both before and after the K subdivisions, I would then multiply each interval's probability by K . The goal is then to come up with a new height, when subdivided, preserves total area. So now the height does not represent the probability of the subinterval but rather some amount that is proportional both to its probability and K , the number of subdivisions. As K increases, the width of each subinterval linearly decreases by $\frac{1}{K}$ factor. So K is inversely proportional to the subinterval width. Like if $K = 3$ the width is $\frac{\Delta x}{3}$ so width decreased by a factor of $\frac{1}{3}$ and again if I multiply K by 3 to $K = 9$ the width decreased by a factor of $\frac{1}{3}$ to $\frac{\Delta x}{9}$. So the height being directly proportional to K means it is inversely proportional to the subinterval width $\frac{\Delta x}{K}$. So the idea is height proportional to prob and inversely proportional to width and that's how I see density.