Relative frequency density, or RFD
(also known as "probability density")

Given a data set X, we define relative frequency density, or RFD, by

(*) RFD (of a range of data values) $= \dfrac{\text{frequency}}{B \cdot n}$

Here:

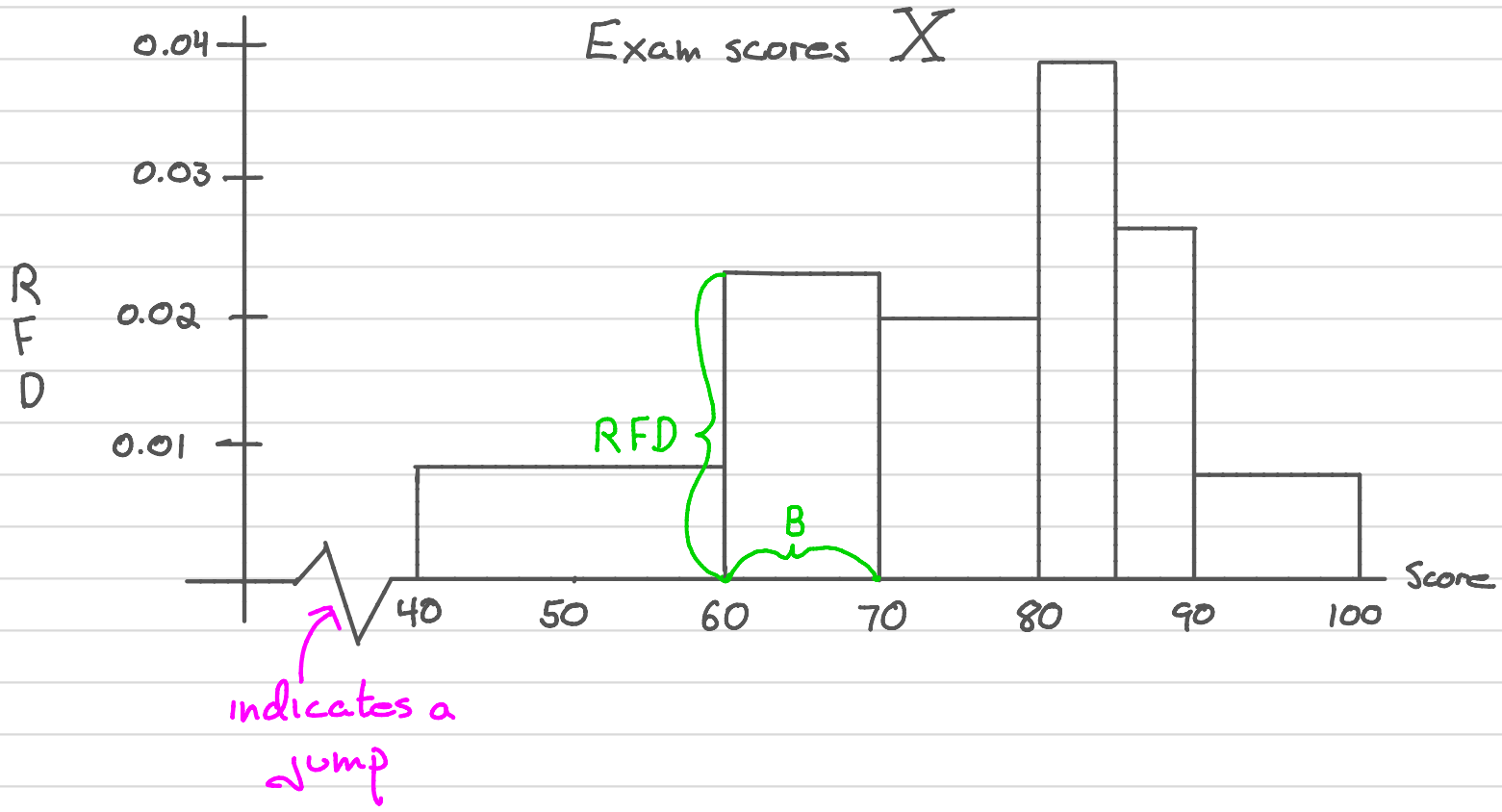
- n is the total number of data points in X.
- B is the length of the range in question. (B stands for bin, meaning range.)
- "frequency" means the number of data points in that range (i.e. bin).

Example: Consider this data set X of 68 exam scores.

Range	Frequency	RFD
$[40, 60)$	11	$11/(20 \cdot 68) = 0.0081$
$[60, 70)$	16	$16/(10 \cdot 68) = 0.0235$
$[70, 80)$	14	$14/(10 \cdot 68) = 0.0206$
$[80, 85)$	13	$13/(5 \cdot 68) = 0.0382$
$[85, 90)$	9	$9/(5 \cdot 68) = 0.0265$
$[90, 100)$	5	$5/(10 \cdot 68) = 0.0074$

(Note that different bins can have different lengths!)

Here's an RFD histogram:



Question: so what? Why do we bother with RFD?

Answer: in an RFD histogram, probability = area!

To see this, recall (*):

$$RFD = \frac{\text{frequency}}{B \cdot n}.$$

Multiply by B:

$$RFD \cdot B = \frac{\text{frequency}}{n}.$$

= area of the bar over a given bin (see RFD histogram above)

$\frac{\text{frequency}}{n}$ = proportion (or percentage, or fraction) of the data that's in the given bin = probability that a data point, selected at random, will be in that bin.

SO: probability is about area (definite integrals, Riemann sums, etc).

Example.
In the above RFD histogram,

$P(60 \le x < 90)$ = sum of areas of bars above $[60, 90)$
$= 0.0235 \cdot 10 + 0.0206 \cdot 10 + 0.0382 \cdot 5 + 0.0265 \cdot 5$
$= 0.7645$
$= 76.45\%$.

this means: the probability that an element x of X, selected at random, lies in $[60, 90)$

General notes on RFD histograms:

- Bins should not overlap.
- All bins, taken together, should cover all data values.
- We can at best approximate $P(a \le x < b)$ if a or b is not the endpoint of a bin.

E.g. for X as above, we might approximate

$P(53 \le x < 66) \approx 7 \cdot 0.0081 + 6 \cdot 0.0235$

length of bin $[53, 60)$ — length of bin $[60, 66)$

$= 0.1977 = 19.77\%$.