How to calculate Z scores in R

Cholesterol levels in men 18-24 years of aage are normally distributed with a *mean of 178* and a *SD of 40*. [1]

1. In what percentile is a man with cholesterol level of 190?

To convert a Z score to a percentile, you need to use the pnorm() function in R to calculate the cumulative probability up to the Z score. This cumulative probability represents the percentile below the Z score, not above it.

- Z = Obs - mean/SD
- Z = `(190-178)/40`

```
[1] 0.3
```

- The Z score of 0.3 is the SD above the mean.
  To convert to a percentile, use `pnorm() * 100`
- Convert Z score to Probability

- **percentile <- pnorm(Zscore)**

```
percentile <- pnorm(0.3)
percentile
```

```
[1] 0.6179114
```

[1] These problems are found in the tutorial from David Longstreet's *statisticsisfun* channel, "How to calculate Z scores" (2011), https://www.youtube.com/watch?v=Ypf6eAP9aFg

2. What is the probability that a man in this age group has a cholesterol level greater than 145?

i.e. All of the are to the right of 145.

```
z145 <- (145-178)/40

1- pnorm(z145)
```

```
[1] 0.7953142
```

3. What is the probability that a man in this age group has a cholesterol level greater than 178?

```
z178 <- (178-178)/40

1-pnorm(z178)
```

```
[1] 0.5
```

4. What Cholesterol level corresponds to the 75th percentile?

$$Z = (X - mean)/SD$$

You would apply the inverse cdf to the percentile ranks to convert them to quantiles, so if you want standard normals, z=Φ−1(p) should do what you seem to be asking for.

- To convert a value, let's call it X, to the 75th percentile in R when the mean is 178 and the standard deviation is 40, you need to calculate the corresponding Z score and then find the value that corresponds to that Z score in a standard normal distribution.
- The formula to calculate the Z score is:
  - Z = (X - mean) / standard deviation
- **To convert the Z score to the corresponding percentile, you can use the *qnorm()* function in R**.
  - The *qnorm()* function calculates the quantile (value) of a standard normal distribution given a probability.

- Z <- qnorm(percentile/100) = Z score corresponding to the 75th percentile, then, `X <- Z & SD + mean` calculates the X corresponding to that Z score.

```
mean <- 178
std_dev <- 40
percentile <- 75

Z <- qnorm(percentile/100)
X <- Z * std_dev + mean
X
```

```
[1] 204.9796
```

5. What is the probability that a man in this age group has a cholesteral level between 155 and 185?

```
probability1 <- pnorm((185-178)/40) - pnorm((155-178)/40)
probability1
```

```
[1] 0.2868145
```

6. If a sample of 100 men is selected, what is the probability that the mean cholesterol level is greater than 175?

$$Z = \frac{X - \text{mean}}{\frac{\text{SD}}{\sqrt{\text{sample size}}}}$$

```
Z100 <- (175-178)/(40/10)

p <- 1-pnorm(Z100)
p
```

```
[1] 0.7733726
```