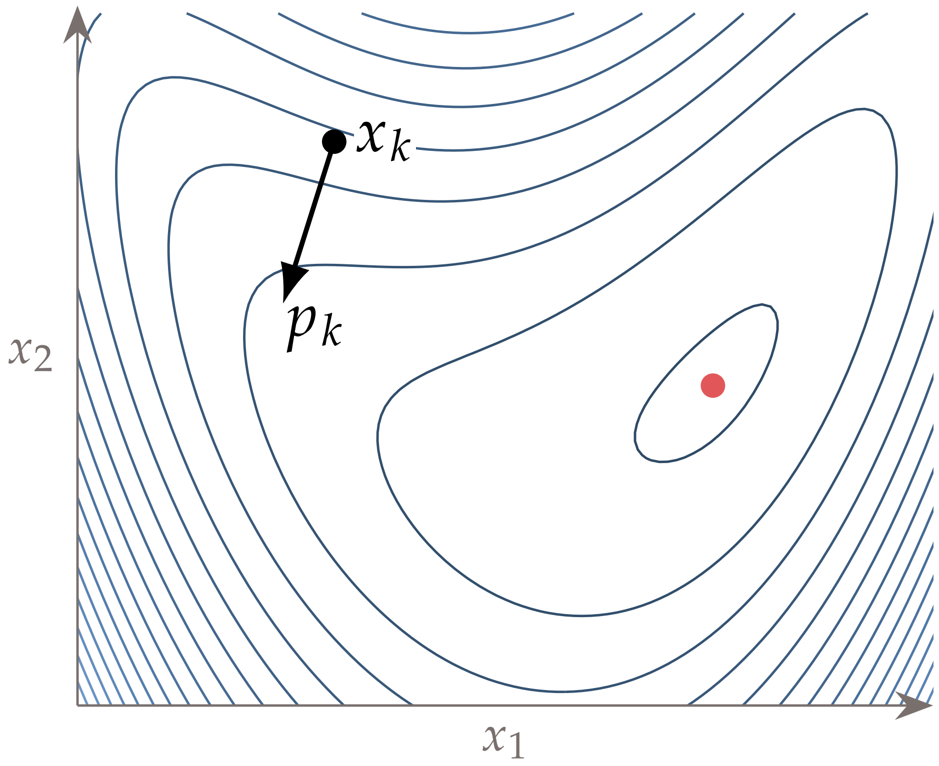
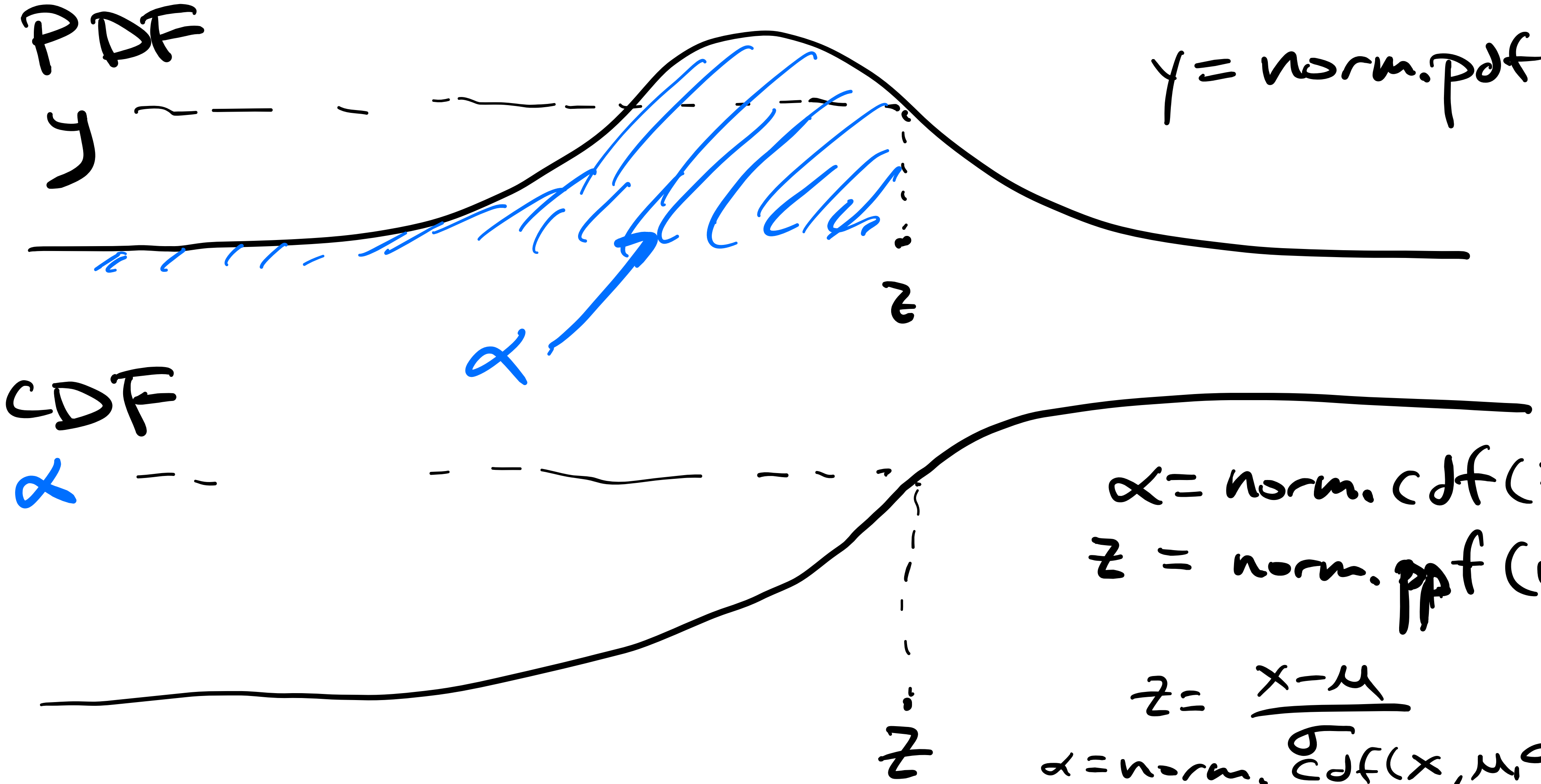
Confidence Intervals



ME EN 275
Andrew Ning
aning@byu.edu

CDF and PPF (inverse CDF)

from scipy.stats import norm



$y = \text{norm.pdf}(z)$

$\alpha = \text{norm.cdf}(z)$

$z = \text{norm.ppf}(\alpha)$

$z = \frac{x-\mu}{\sigma}$

$\alpha = \text{norm.cdf}(x, \mu, \sigma)$

Example

scipy.stats.norm. cdf
ppf

Battery lifetimes for a given application are normally distributed with mean 50 hrs and stdev = 5 hrs.

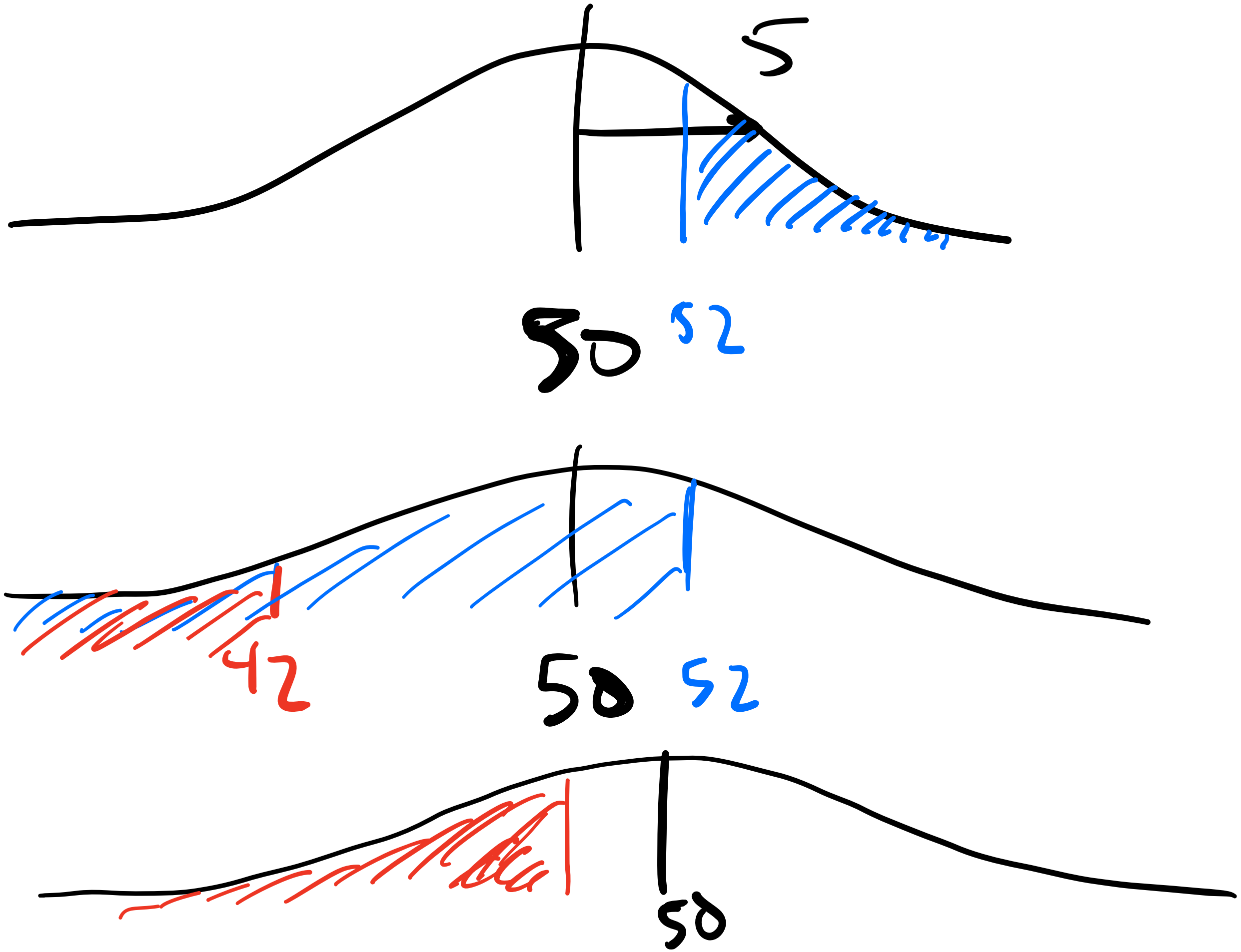
What is probability that a random battery lasts longer than 52 hours?

What is probability that a random battery lasts between 42 and 52 hours?

What is the 40th percentile of battery lifetimes?

Statistics for Engineers and Scientists, Navidi

Example

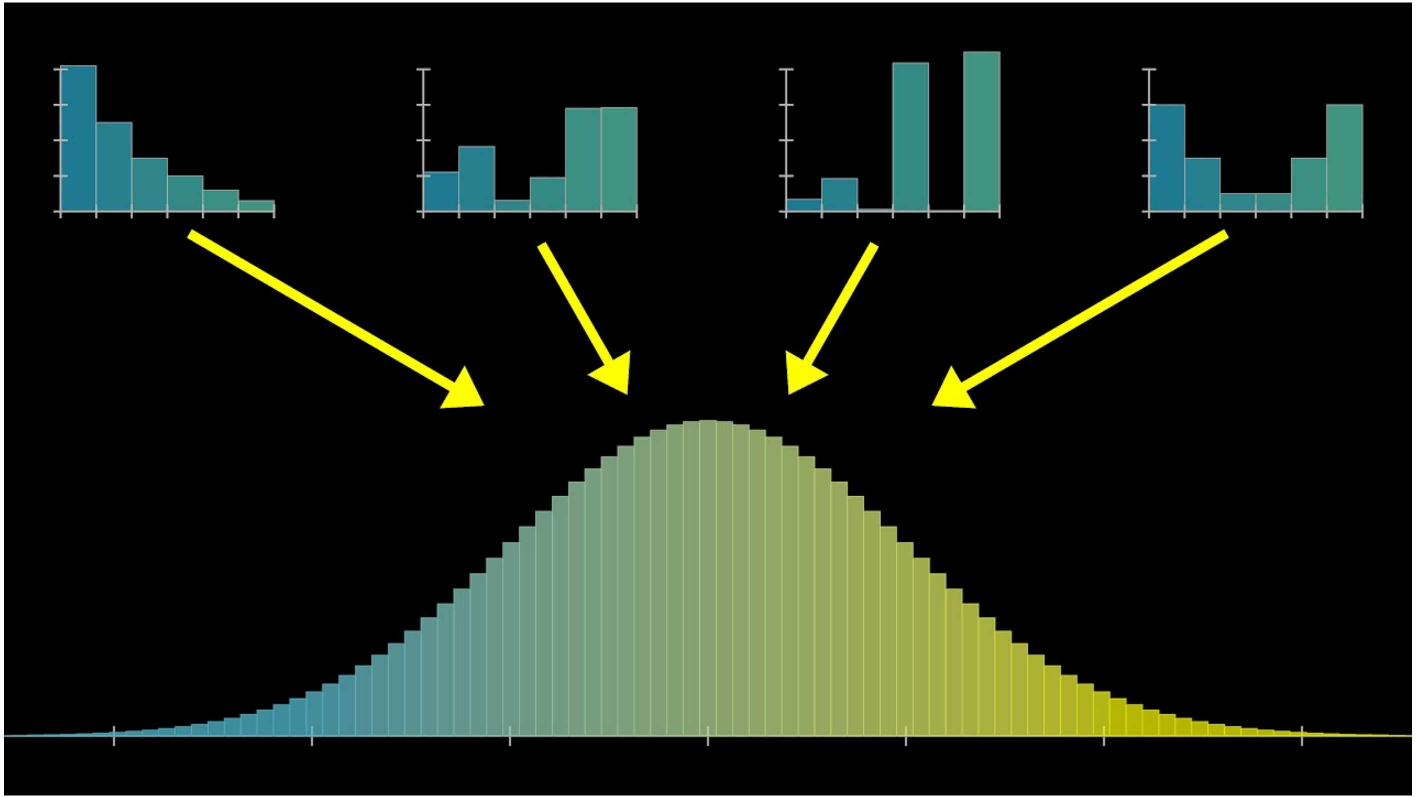


Other Common Distributions

https://en.wikipedia.org/wiki/List_of_probability_distributions

Motivation

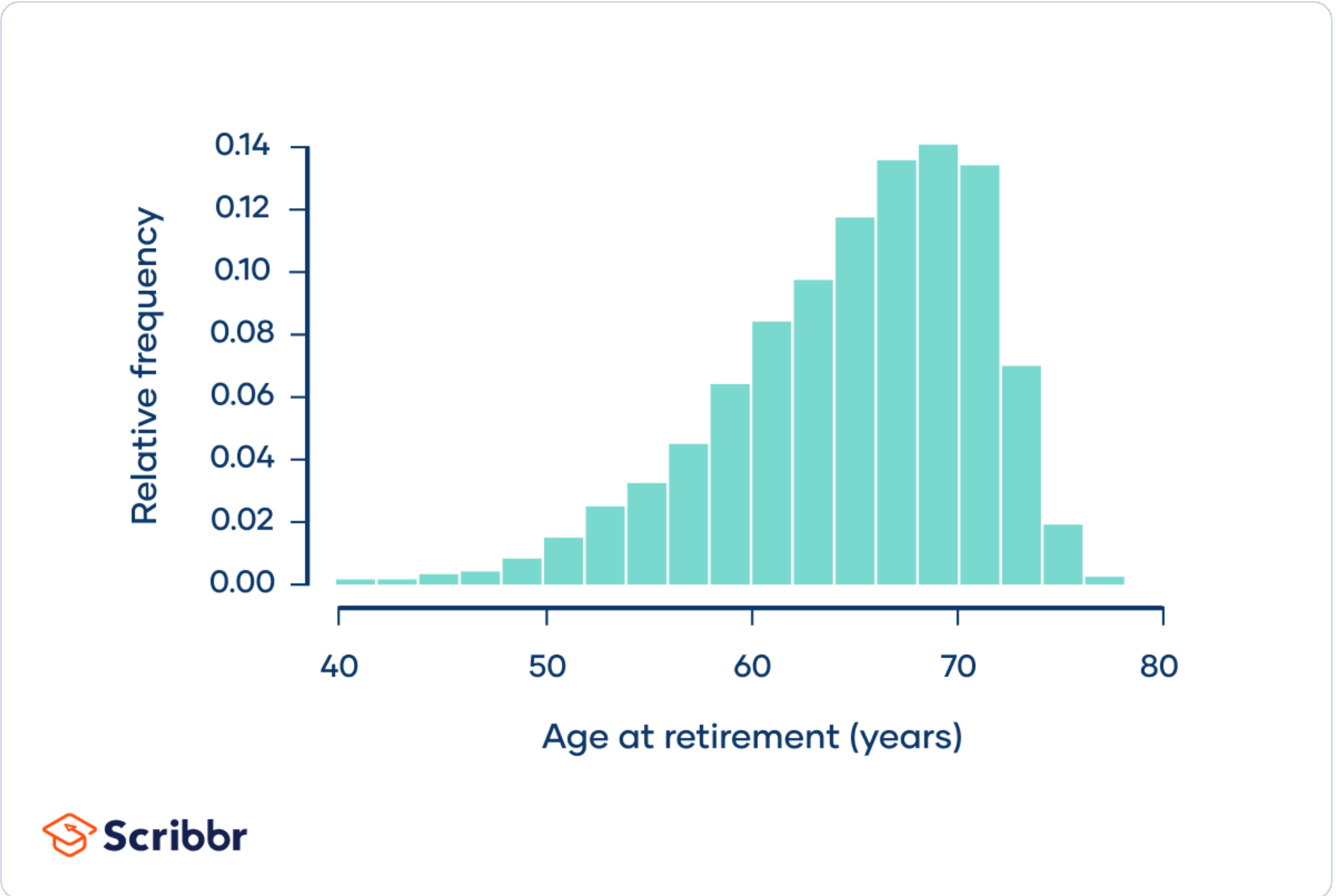
Suppose I sampled the height of 100 people in Provo, UT computed their mean height. I would expect the mean I measured to not be quite the true mean of all people in Provo - perhaps higher or lower. But how good is my estimate? Can I provide a reasonable range for what the mean height is?



Central Limit Theorem

sampling distribution of the mean will always be normally distributed (assuming sample is large enough, say > 30)

Example (from scribbr.com)



68	73	70	62	63

mean = (68 + 73 + 70 + 62 + 63) / 5

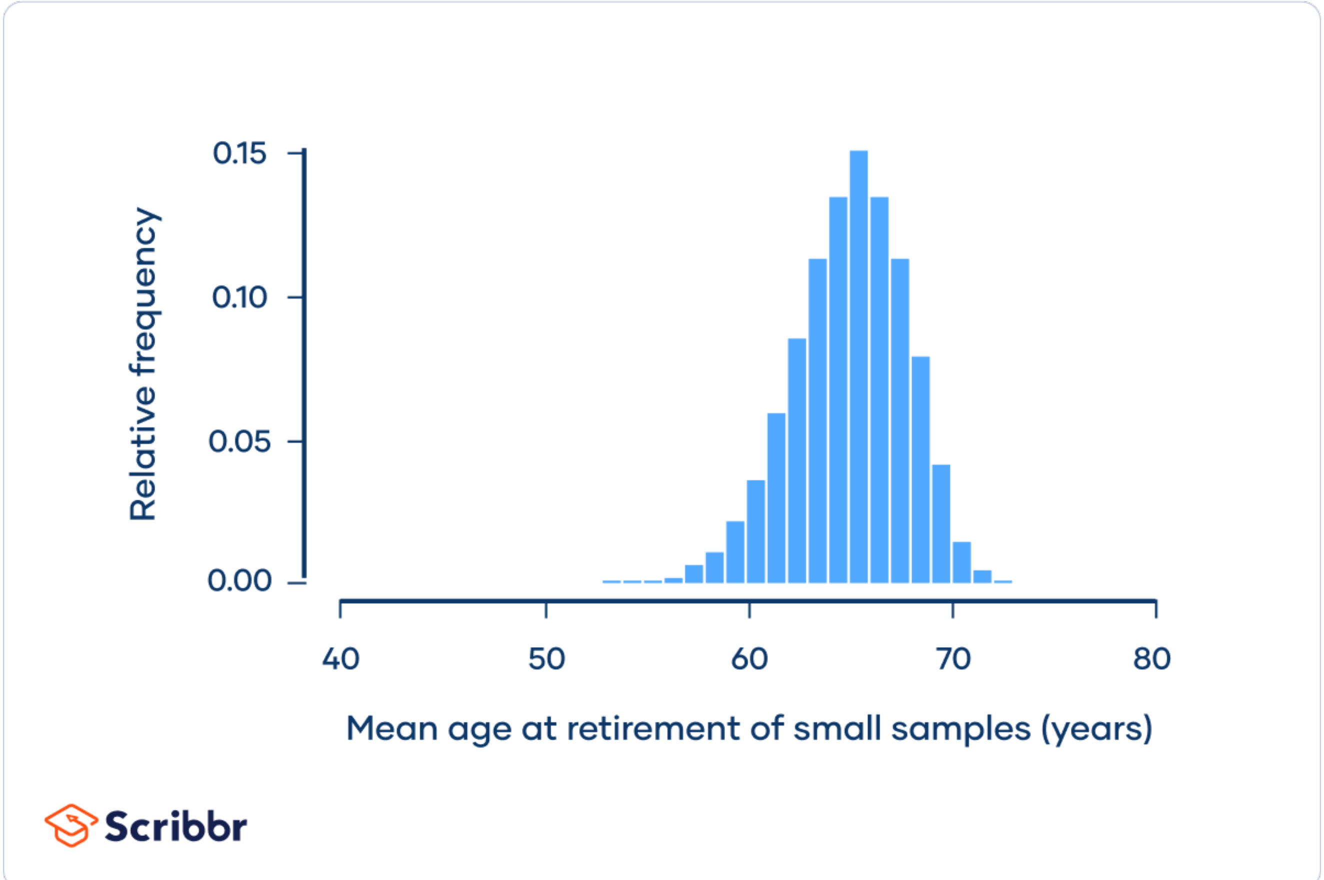
mean = 67.2 years

sampling distribution of the mean

60.8	57.8	62.2	68.6	67.4	67.8	68.3	65.6	66.5	62.1

sampling distribution of the mean

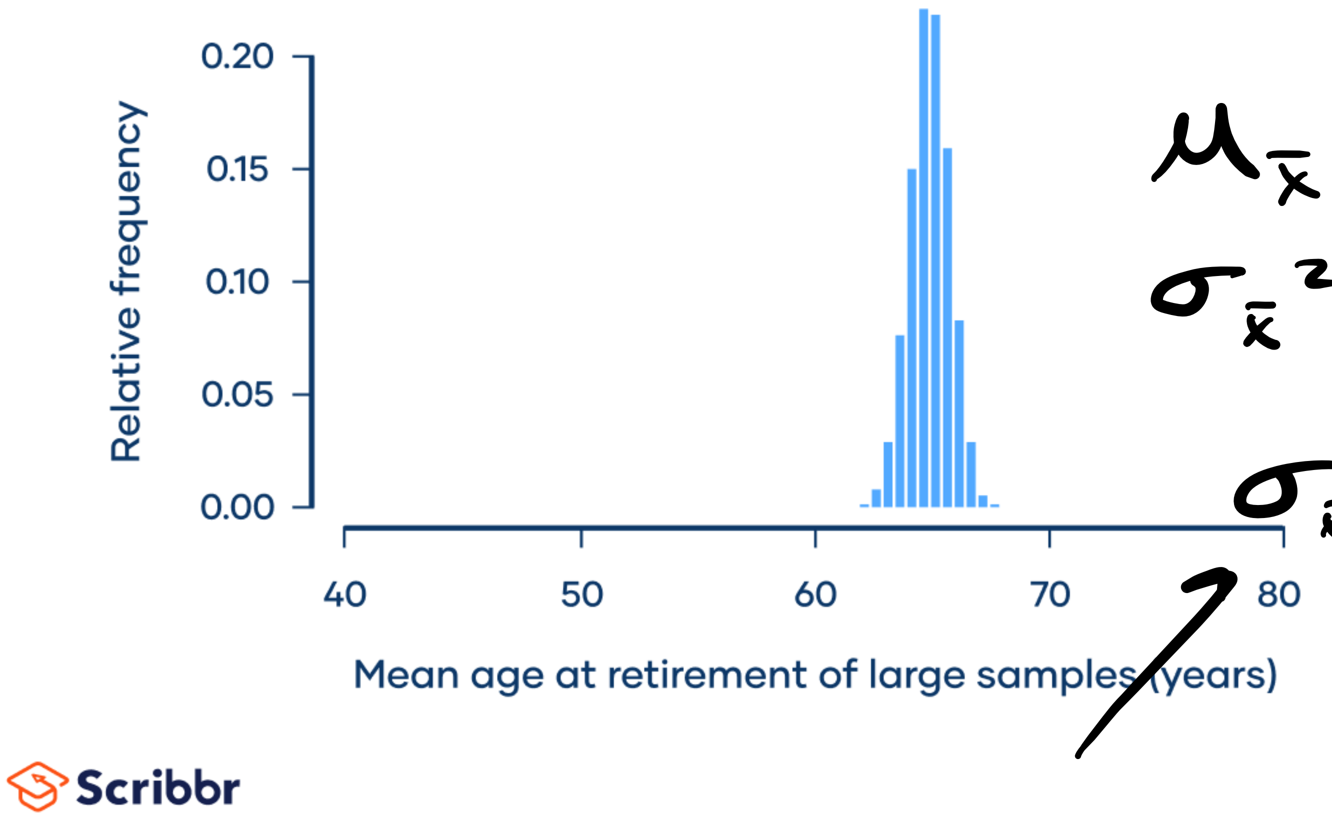
60.8	57.8	62.2	68.6	67.4	67.8	68.3	65.6	66.5	62.1



73	49	62	68	72	71	65	60	69	61
62	75	66	63	66	68	76	68	54	74
68	60	72	63	57	64	65	59	72	52
52	72	69	62	68	64	60	65	53	69
59	68	67	71	69	70	52	62	64	68

mean = 64.8 years

$$\bar{x} = \frac{x_1 + x_2 + \dots x_n}{n}$$



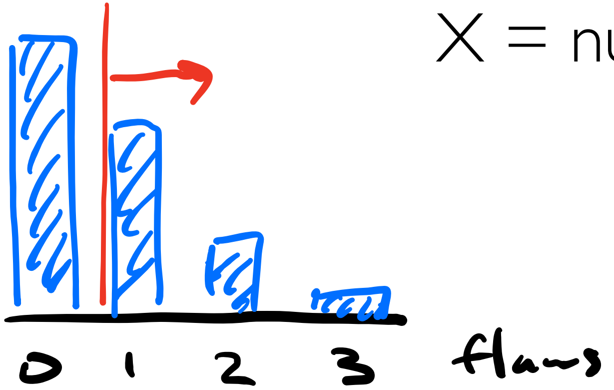
$$\mu_{\bar{x}} = \mu_x$$

$$\sigma_{\bar{x}}^2 = \frac{\sigma_x^2}{n}$$

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

standard error = stdev of sample mean

Example



X = number of flaws in a copper wire

$\mu_{\bar{x}} = \mu_x$

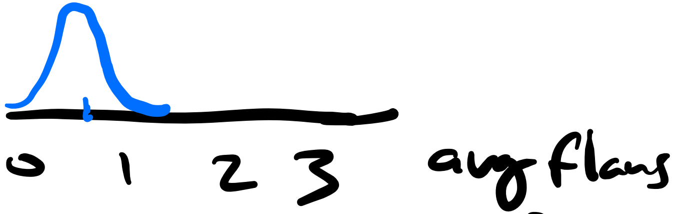
$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$

x	$P(X = x)$
0	0.48
1	0.39
2	0.12
3	0.01

$\mu_x = 0(0.48) + (1)(0.39) + \cdots$
$= 0.66$

$\mu_{\bar{x}} = 0.66$

$\sigma_{\bar{x}} = \sqrt{.5244}/\sqrt{100}$



from samples of n = 100

If I sample 100 of them, what is probability that they will have an average number of flaws less than 0.5?

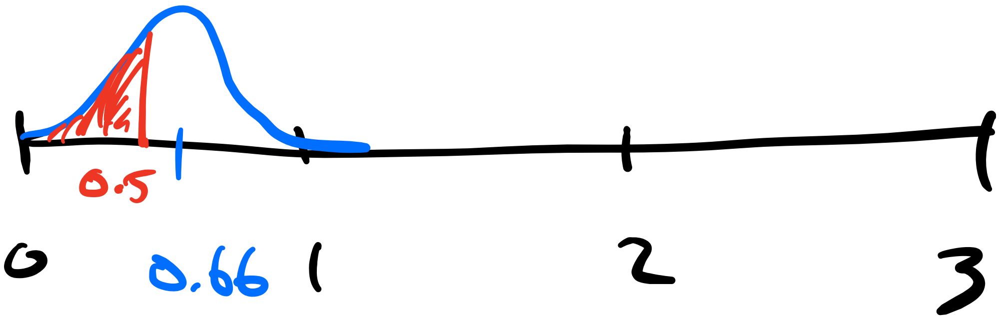
$\sigma_x^2 = (0-\mu_x)^2(0.48) + (1-\mu_x)^2(0.39) + \cdots$
$= 0.5244$

Statistics for Engineers and Scientists, Navidi

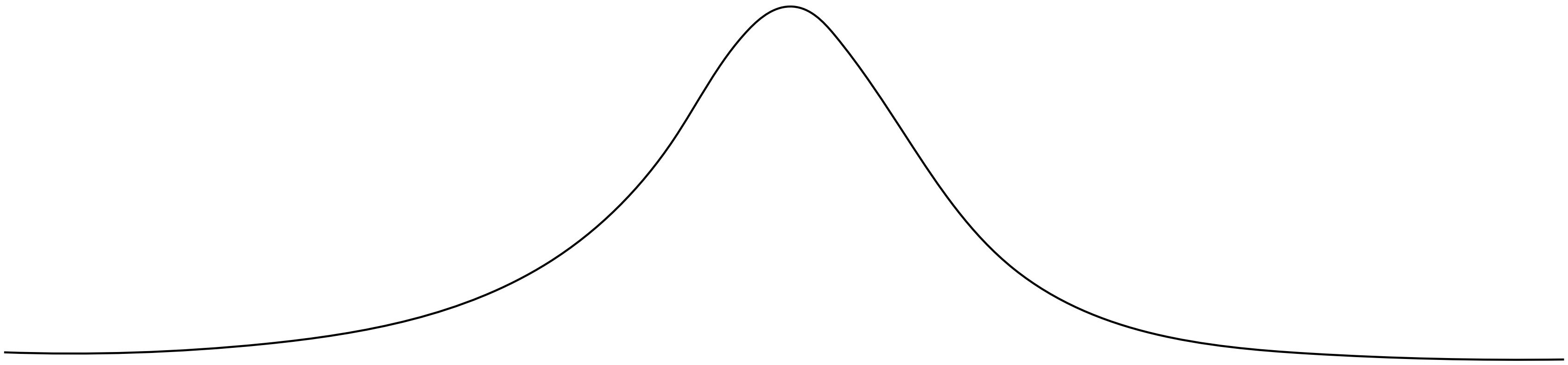
Example

$$\text{norm.cdf}\left(0.5, 0.66, \sqrt{\frac{.5244}{100}}\right)$$

$\sim 1.35\%$



Example (heights)



Other confidence intervals

Example

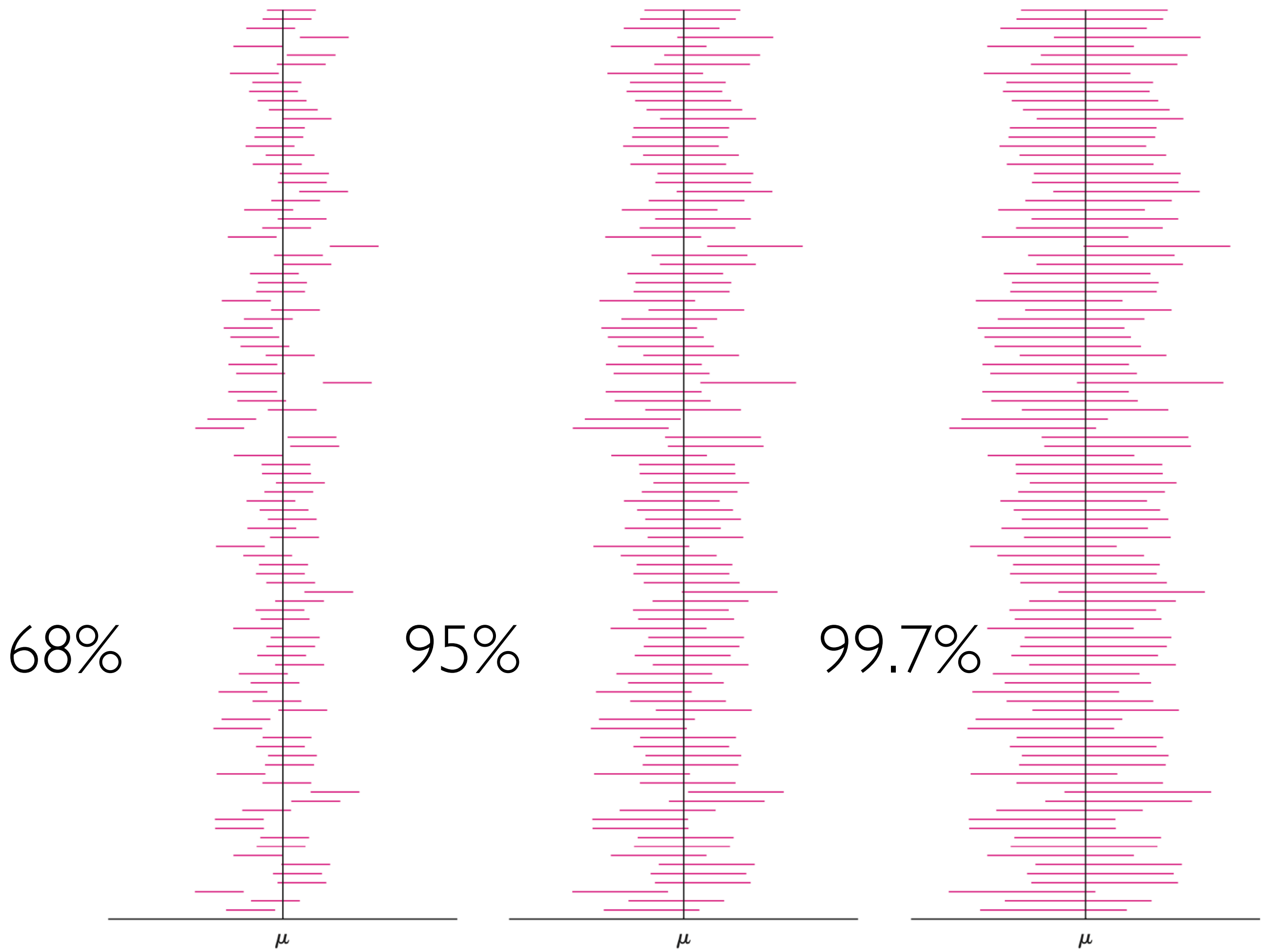
sample size = 50

$\bar{x} = 12$

$s_x = 1$

find 80% confidence interval

What confidence level should I use?



Statistics for Engineers and Scientists, Navidi