1. (page 310, exercise 6) A zircon semiconductor is critical to the operation of a superconductor and must be immediately replaced upon failure. Its expected lifetime is 100 hours, and its standard deviation is 34 hours. If 22 of these semiconductors are available, approximate the probability that the superconductor can operate for the next 2000 hours. (That is, approximate the probability that the sum of the 22 lifetimes exceeds 2000.)

Xi = lifetime of the i-th zircon semiconductor

E(X) = 100 h ; SD(X) = 34 h ; sample size n = 22

We know that $\sum_{i=1}^{22} X_i$ has approximately a normal distribution (if n is large enough), with:

Mean = (sample size) . E(X) = 22 x 100 = 2200

Var = sqrt(sample size) . Var(X) = sqrt(22) . 34^2 = 5422.121

=> SD = sqrt(5422.121) = 73.64

$P\left(\sum_{i=1}^{22} X_i > 2000\right) =?$

Since this is normally distributed, we can calculate through the Z value:

= $P\left(\frac{\sum_{i=1}^{22} X_i - 2200}{73.64} > \frac{2000 - 2200}{73.64}\right)$ = P(Z > -2.71592) ≈ 1 – P(Z ≤ -2.72) = 1 – ϕ(-2.72) = 1 – 0.00426 = **0.99674 = 99.674%**

2. (Page 311, problem 7) The amount of paper a print shop uses per job has mean 200 pages and standard deviation 50 pages. There are 2300 sheets of paper on hand and 10 jobs that need to be filled. What is the approximate probability that 10 jobs can be filled with the paper on hand?

X = number of pages used per job

E(X) = 200 ; SD(X) = 50

n = 10 jobs to do

We know that $\sum_{i=1}^{10} X_i$ has approximately a normal distribution (if n is large enough), with:

Mean = (sample size) . E(X) = 10 x 200 = 2000

Var = sqrt(sample size) . Var(X) = sqrt(10) . 50^2

=> SD = 50 . sqrt(10) = 158.11

$P\left(\sum_{i=1}^{10} X_i \leq 2300\right) =?$

Since this is normally distributed, we can calculate through the Z value:

= $P\left(\frac{\sum_{i=1}^{10} X_i - 2000}{158.11} \leq \frac{2300 - 2200}{158.11}\right)$ = P(Z ≤ 1.897) = ϕ(1.897) ≈ ϕ(1.90) = 0.97128 = 97.128%

3. The average amount of coffee made in a certain meeting room is 15 liters per day with the variance of 9 liters. The amount of coffee made is monitored on 10 random days in this certain meeting room. Approximate what is the probability that the average of coffee made during these days is

(a) more than 20 liters?

(b) between 7 and 11 liters?

X = amount of coffee drank per day in the meeting room

E(X) = 15 ; Var(X) = 9 => SD(X) = sqrt(9) = 3

n = sample of days = 10

Average X = $\bar{X} = \frac{1}{10}\sum_{1=1}^{10} X_i$ has approximately a normal distribution (if n is large enough) with:

Mean = E(X) = 15

SD = SD(X) / sqrt(10) = 3/sqrt(10) = 0.949

----

a) P($\bar{X}$ > 20) = ?

= $P\left(\frac{\bar{X_i}-15}{0.949} > \frac{20-15}{0.949}\right)$ = P(Z > 5.27) = 1 – P(Z ≤ 5.27)

The value P(Z ≤ 5.27) will be very close to 1 (=100%), therefore:

P(X > 20) = 1 – P(Z ≤ 5.27) ≈ 1 – 1 = 0

----

b) P(7 < $\bar{X}$ < 11) = ?

= P($\bar{X}$ < 11) – P($\bar{X}$ ≤ 7) = P[Z < (11-15)/0.949] – P [Z ≤ (7 – 15)/0.949] = ϕ(-4.215) – ϕ(-8.430) ≈ 0.00 – 0.00 = 0.00

4. (Page 342, problem 12) Suppose that the systolic blood pressure of a worker in the mining industry is normally distributed. Suppose also that a random sample of 13 such workers yielded the following systolic blood pressures:

129, 134, 142, 114, 120, 116, 133, 142, 138, 148, 129, 133, 141.

(a) Estimate the mean systolic blood pressure of all miners.

(b) Estimate the standard deviation of the systolic blood pressure.

(c) Use the estimates in parts (a) and (b) along with the fact that the blood pressures are normally distributed to obtain an estimate of the proportion of all miners whose blood pressure exceeds 150.

X = systolic blood pressure of a worker ; n = number of workers sampled = 13

a) Mean(A) = $\frac{1}{n}\sum_{1=1}^{n} X_i$ = (129 + 134 + … + 133 + 141)/13 = 1719/13 = 132.23

b) Var(X) = E { [X – E(X)]^2 } = [(129 – 132.23)^2 + (134 – 132.23)^2 + … + (141 – 132.23)^2] / 13^2 =

= 7.8125

SD = sqrt(Var) => SD = sqrt(7.8125) = 2.795

c) $P(X > 150) = ?$

$= P(Z > (150 – 132.23)/2.795) = 1 – P(Z \leq 6.3578) \approx 1 – 1 = 0.00$

There are approximately 0.00% of workers with systolic blood pressure above 150.

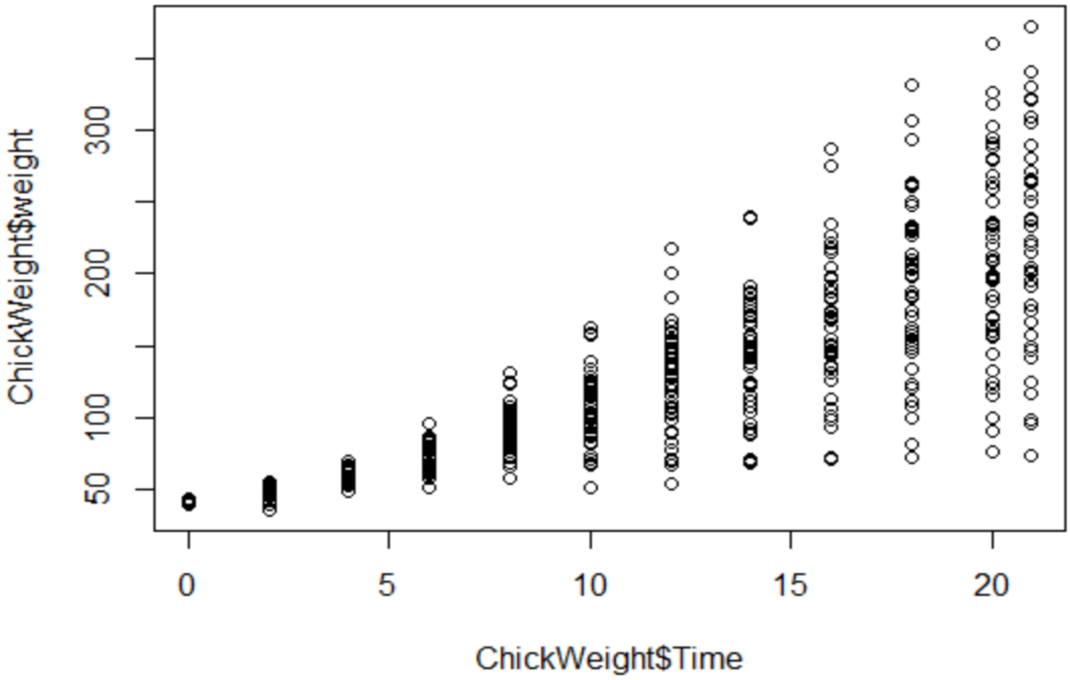
5. Week 2 exercises featured the Bayes' theorem. Ross's book focuses on opening up the classical content of statistics, the so-called *frequentist statistics* and we learn about so-called *Bayesian statistics* in other ways (as well as in Part II of the course and also through Ross's book; The 4th edition has a some amount of Bayesian statistics). Bayes' theorem dates back to the 18th century. As late as the 19th century, statistical inference really relied upon to Bayesian way of thinking. However, from the beginning of the 20th century for about 80 years the Bayesian inference was mostly in the dark. The last twenty over the years, however, Bayesian statistics has become mainstream in many applications.

   Article in Significance (a journal popularizing statistics) *Are our brains Bayesian?* from 2016 can be found from `https://rss.onlinelibrary.wiley.com/doi/full/10.1111/j.1740-9713.2016.00935.x`. Try to find an explanation for the title of the article from this article. It is enough to answer in a few sentences. You don't have to understand the article completely, not even close! An important skill the world of science is that you just look at or browse scientific texts. Often even a working scientist understands only part of the literature she reads, unless it is exactly from her own area of expertise. However, we need to keep up the pace and follow along all kinds of ideas.

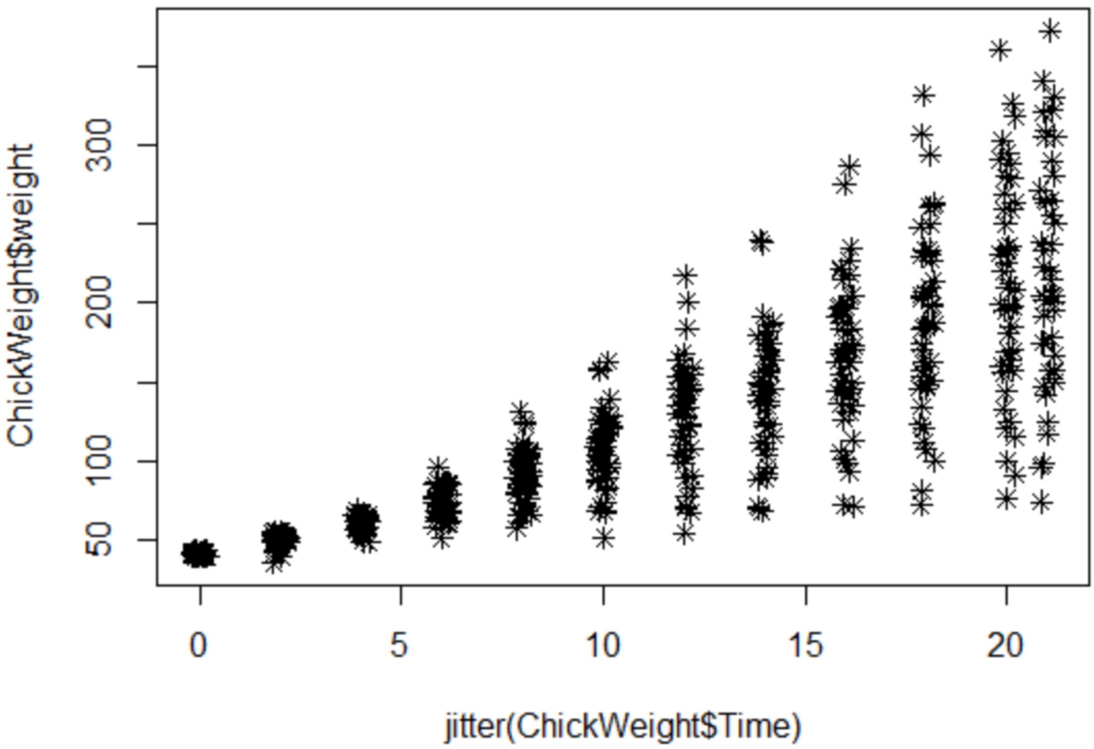
The article is entitled " Are our brains Bayesian?" because AI researchers have been using algorithms which apply the principles of Bayesian inference in trying to replicate the way our brains work.

(R plots in the next page)

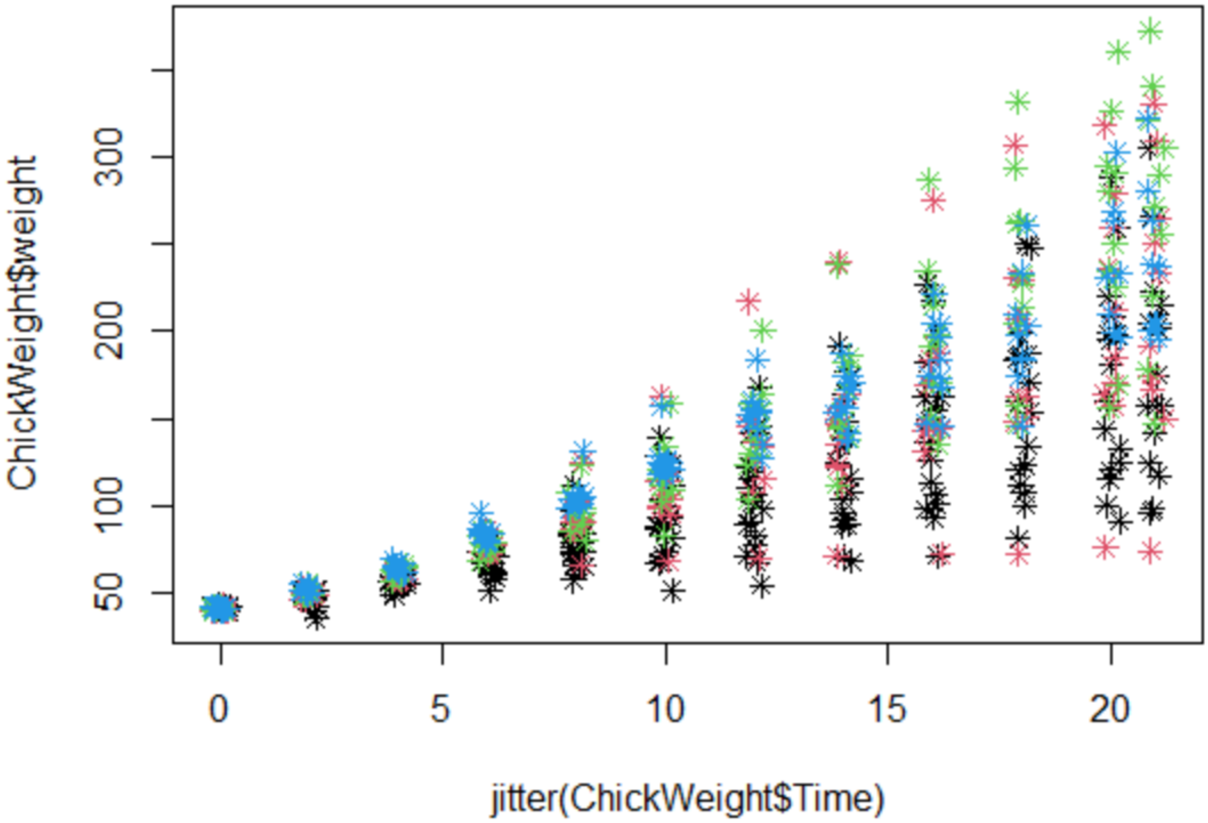
3 a) Scatter plot of Weight (vertical axis) versus Time (horizontal axis)



3 b) the points in the plot are now stars, which have been jittered with the jitter() variable



3 c) Colouring the data points by different Diet



3 d) Adding a legend so that we know what colour corresponds to which Diet (diet 1, 2, 3 or 4)