Some sampling distribution problems

Sampling distribution of the mean of normally distributed data. (*These problems are not exactly the same as those worked through in class/lab.*)

1. I eat a breakfast cereal every day. The amount of saturated fat in a serving is normally distributed with mean 25 g and standard deviation 4 g. Find the probability that my saturated fat intake on any day is below 27 g.

Solution: This is a regular old normal probability calculation. We turn this value into a Z score getting Z=(27-25)/4 = 2/4 = 0.5. We want the area to the left; this is 0.6915.

2. If I eat it every day for a week, find the sampling distribution of the daily average fat intake over the week. Use it to find the probability that the average daily saturated fat intake for the week was less than 27 g

Now we are into the sampling distribution of the mean. A week has seven days, so the average fat content has distribution N(25, $4/\sqrt{7}$), or N(25,1.51). The probability that the mean is below 27 g requires conversion to a Z score: (27-25)/1.51 = 1.32. The probability a Z below 1.32 is 0.9066.

3. If I eat it for a 30-day month, what is the probability that the average daily saturated fat intake for the month was less than 27 g.

Now we are up to n=30. and so the monthly mean has distribution N(25,$4/\sqrt{30}$) = N(25,0.73). To find the probability asked for, we compute Z=(27-25)/0.73 = 2.74. The area to the left of 2.72 under the standard normal is 0.9969.

These three examples show how increasing sample size leads to a sample mean that tends to be closer and closer to the true mean μ. Although on any given day there is quite a high probability of fat intake over 27 g, once you aggregate over several days this probability goes down dramatically.

4. What is the probability that my total saturated fat intake for the month was less than 810 g?

The month is 30 days long. So for the total for the month to be below 810 g, we would need the daily average to be below 810/30 = 27 g. We just figured this probability as 0.9969.

5. What is the probability that it was less than780 g?

For this to happen, the daily average would have to be below 780/30 = 26 g. This gives a Z score of Z=(26-25)/0.73 = 1.37. The probability of this happening is 0.9147.

Central limit theorem

The weight of adult males has a mean of around 65 kg and a standard deviation of 20 kg. The distribution is not normal.

6. Can you calculate the probability that a randomly selected individual weighs more than 75 kg?

No. If you don't know the distribution, you can not compute the probability, even though you know μ and σ.

7. Suppose that a sample of size 16 is big enough for the central limit theorem to apply to the average weight of a random sample of adult males. What is the probability that the average weight of 16 randomly selected males will exceed 75 kg?

We know that if individual values have a population mean of μ and a population standard deviation of σ, the mean of a sample if size n will have a population mean of μ and a standard deviation of $\sigma/\sqrt{n}$. On its own, this does not help us very much. But we are also told that n=16 is large enough for the central limit theorem to hold. So we can use the result that the mean weight of these 16 random individuals is N(65, 20/√16), or N(65,5). To find the probability that the average is below 75 kg, we need to convert 75 kg into a Z score by getting Z=(75-65)/5 = 2.0 We want the area to the right of this; this is 1-0.9772 = 0.0228.

8. Find the 90% range within which the average weight of a random sample of 16 adult males will lie?

To get the 90% range we start with the Z values defining the central 90% of the standard normal. These values are Z=-1.64 and Z=1.64. (Review the earlier material on working with the normal if this part is not clear to you.) Next, we need to convert these Z scores back to the kg scale for the average body weight of the 16 individuals. This gives the two limits as 65-1.64*5 = 56.8 kg for the lower end, and 65+1.64*5 = 73.2 kg for the upper end.

9. Redo question 8 for a random sample of 100 adult males.

All that changes here is the standard deviation of the sampling distribution of the sample mean. For a sample of size n=100, this is 20/√100 = 20/10 = 2 kg. So now when we convert the two Z scores back to the kg scale, they give 65-1.64*2 = 61.72 for the lower limit, and 68.28 for the upper end.

The amount of baggage a plane passenger checks is random, with a mean of 20 lbs and a standard deviation of 30 pounds

10. Can you calculate the probability that a passenger checks more than 30 pounds of baggage?

No. Look at the mean and standard deviation. The 68/95/almost all rule tells you that for a normal distribution about 16% of data are below $\mu-\sigma$ But with these numbers, $\mu-\sigma$=20-30 = -10kg. No one has baggage with a negative weight, so clearly the amount of checked baggage cannot possibly follow a normal distribution. We are back in the same area as the last problem where we did not know the distribution of an individual X.

11. A plane carries 100 passengers. It can handle 3000 pounds of checked baggage, that is, a maximum of 30 pounds per passenger. What is the probability that a random load of 100 passengers will check too much baggage for the plane to handle?

The sample size of 100 is probably large enough for the central limit theorem to work. So the average amount of baggage per passenger will be close to N(20,30/√100) = N(20,30/10) = N(20,3). The plane is in trouble if this average exceeds 30 pounds. To find the probability of this we convert 30 pounds to a Z score, getting Z=(30-20)/3 = 10/3 = 3.33. The probability to the right of 3.33 is 1-0.9996 = 0.0004, so the risk is actually pretty slight.