Problem 2.1. (9 points) Source: “Probability” by Jim Pitman.

Consider the sample average $\bar{X}_n$ of $n$ independent random variables, each of them uniformly distributed on $[0, 1]$. Find the smallest sample size $n$ such that $P[\bar{X}_n < 0.51]$ is at least 90%. Use the Central Limit Theorem.

Solution: From the CLT we know that approximately

$$\bar{X}_n \sim N(mean = 1/2, var = 1/(12n)),$$

i.e.,

$$\frac{\bar{X}_n - 0.5}{\sqrt{1/12n}} \sim N(mean = 0, var = 1).$$

So, with $Z \sim N(0, 1)$, the probability we are looking for is

$$P[\bar{X}_n < 0.51] \approx P[Z < \frac{0.51 - 0.5}{\frac{1}{\sqrt{12n}}}] = \Phi(0.02\sqrt{3n})$$

Consulting the standard normal tables, we get that the sufficient condition for the above probability to be at least 0.90 is

$$0.02\sqrt{3n} \geq 1.29 \Rightarrow 3n \geq 4160.25.$$ 

Our final answer is $n \geq 1387$.

Problem 2.2. Source: “Probability” by Pitman.

Suppose that on average 1 in 100 people gave a certain genetic marker.

(i) (3 points) Suppose a simple random sample of 50 is selected and tested. What is the probability that at least one of them will have the genetic marker?

Solution: The number of people in the random sample who have the genetic marker $Y$ has the binomial distribution with the number of trials equal to 50 and with probability of success in every trial equal to 1/100.

$$P[Y \geq 1] = 1 - P[Y = 0] = 1 - (99/100)^{50} = 0.395.$$ 

(ii) (3 points) About how many people would have to be tested so that the probability of having at least one positive result would be at least 99%? If this number of people were tested, what is the expected number of positives?

Solution:

$$0.99 \geq 1 - 0.99^n \quad \Leftrightarrow \quad 0.99^n \geq 0.01 \quad \Leftrightarrow \quad n \geq \frac{\ln(0.01)}{\ln(0.99)} = 458.11$$

Hence, $n \geq 459$. The expected number of positive results is $459 \times 0.01 = 4.59$. 

Instructor: Milica Ćudina