Limit Theorem Clinic

Lesson 75 · Module 5 — Limit Theorems · Probability & Statistics · dailymaths.uk/probstat

Practice problems

1. A fair six-sided die is rolled $n = 100$ times. Use the Central Limit Theorem to estimate the probability that the sample mean $\bar{X}$ is between 3.2 and 3.8.

2. You are given a random variable Y with mean $\mu = 50$ and standard deviation $\sigma = 10$. Without knowing the distribution of Y, what is the maximum possible probability that Y is outside the interval $[20, 80]$?

3. A complex system fails if the sum of three independent components' lifetimes, each with a mean of 100 hours and variance of 400, is less than 250 hours. Why is the Central Limit Theorem a poor choice here, and how would a Monte Carlo simulation help?

Solutions

1. A fair six-sided die is rolled $n = 100$ times. Use the Central Limit Theorem to estimate the probability that the sample mean $\bar{X}$ is between 3.2 and 3.8.

   - The mean of a single die is $\mu = 3.5$ and the variance is $\sigma^2 = 35/12 \approx 2.0833$.
   - The standard deviation of the mean is $\sigma_{\bar{X}} = \sqrt{2.0833/100} \approx 0.1443$.
   - Calculate z-scores: $z_1 = (3.2 - 3.5)/0.1443 \approx -2.08$ and $z_2 = (3.8 - 3.5)/0.1443 \approx 2.08$.
   - Find the area between these z-scores: $P(-2.08 < Z < 2.08) \approx 0.9812 - 0.0188 = 0.9624$.

   Answer: 0.9624

2. You are given a random variable Y with mean $\mu = 50$ and standard deviation $\sigma = 10$. Without knowing the distribution of Y, what is the maximum possible probability that Y is outside the interval $[20, 80]$?

   - The interval $[20, 80]$ is centered at the mean 50.
   - The distance from the mean to the boundary is $80 - 50 = 30$.
   - Express this distance in terms of standard deviations: $k = 30/10 = 3$.
   - Apply Chebyshev's inequality: $P(|Y - 50| \geq 30) \leq 1/3^2 = 1/9$.

   Answer: $1/9$

3. A complex system fails if the sum of three independent components' lifetimes, each with a mean of 100 hours and variance of 400, is less than 250 hours. Why is the Central Limit Theorem a poor choice here, and how would a Monte Carlo simulation help?

   - The sample size $n = 3$ is very small, which violates the general rule of thumb ($n \geq 30$) for the CLT to be reliable unless the components are already normally distributed.
   - A Monte Carlo simulation would allow us to generate thousands of triplets of lifetimes based on the actual distribution of the components.
   - By summing these triplets and counting the proportion that fall below 250, we get an empirical probability that doesn't rely on the normal approximation.

   Answer: n=3 is too small for CLT; simulation provides empirical evidence without distribution assumptions.