Summer/Winter/Semester2 | **Tutorial 9** | 2015

This tutorial explores sums of non normal random variables (CLT).

Central Limit Theorem (CLT)
Given a sequence of iid random variables $X_i \sim (\mu, \sigma^2)$ for $i = 1, \ldots, n$.
where $\sigma^2 < \infty$ and n is large,
then the distribution function of $\dfrac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n\sigma^2}}$ tends to the standard Normal.

Less formally, $\displaystyle\sum_{i=1}^n X_i \to N(n\mu, n\sigma^2)$ and $\bar{X} \to N(\mu, \dfrac{\sigma^2}{n})$.

Normal Approximation to the Binomial (CLT special case)
For large n, $X \sim Bin(n, p) \to N(np, np(1-p))$.

Guide: Use when $n > 25$, $np > 5$, $n(1-p) > 5$.

Continuity correction
Given a discrete integer valued RV $X \sim (\mu, \sigma^2)$ and the approximating Normal $Y \sim N(\mu, \sigma^2)$
we adjust by 1/2 to (usually) improve the approximation.
$P(X \geq x) \to P(Y \geq x - 1/2)$
$P(X \leq x) \to P(Y \leq x + 1/2)$

1. Central Limit Theorem

(a) In your own words, explain the Central Limit Theorem.

(b) Why does the Central Limit Theorem apply to the Binomial Distribution?

2. Central Limit Theorem

A new type of electronics flash for cameras will last an average of 5000 hours with a standard deviation of 400 hours. A company quality control engineer intends to select a random sample of 100 of these flashes and use them until they fail.

(a) What is the approximate probability that the mean life time of 100 flashes will be less than 4940 hours?

(b) What is the approximate probability that the mean life time of 100 flashes will be between 4960 and 5040 hours (ie. within 40 hours of μ, the population mean)

3. Normal approximation to Discrete RV

Let X_1, X_2, X_3 be independent rolls of a fair 6-sided die and let $S = X_1 + X_2 + X_3$.

(a) Show that $E(X_i) = 3.5$ and $\text{Var}(X_i) = \dfrac{35}{12}$.

(b) Compute a normal approximation with continuity correction for $P(S \leq 6)$.

(c) Using probability, compute $P(S \leq 6)$ exactly. What is the relative error of the approximation in (b)?

4. Normal Approximation to Binomial

Suppose that $X \sim B(20, 0.25)$.

(a) Write down $E(X)$ and $Var(X)$.

(b) Compute a normal approximation to $P(X \geq 5)$ using a continuity correction.

(c) Use R to compute the exact probability. What is the relative error of the approximation in (b).

(d) Repeat (a)-(c) for $X \sim B(20, 0.1)$.

(e) Why is the relative error in (d) so poor compared to (c)?

5. Normal approximation to Binomial

The MATH1005 exam includes 25 multiple choice questions, each with 5 options. A student randomly guesses each answer independently.

(a) Using a normal approximation with continuity correction, approximate the probability that the student gets at least 8 questions correct.

(b) Use R to compute the exact probability and calculate the relative error of the normal approximation.

6. Poor Normal approximation to Binomial

(a) Find the relative error of the Normal approximation to $P(X \leq 2)$, when $X \sim Bin(25, 0.25)$.

(b) Can you explain why this is a poor appoximation?