Central Limit Theorem

- Under a wide variety of conditions, the sum (and therefore also the mean) of a large enough number of independent random variables is approximately normal (Gaussian).
- Special case: the sum(and therefore also the mean) of independent normal random variables is normal.

Consequences of CLT

- Sampling distributions of means are approximately normal. This is the basis for statistical inference for means, even from non-normal populations.
- Additive processes may be modeled by normal distributions: if a random variable represents the total effect of a sum of many independent small effects, the CLT suggests a normal distribution.

Examples of Additive Processes

- The height of a human (or wingspan of a bird, etc.): many small bones, etc. make up the height
- Errors in sums of rounded numbers.
- Total claims for an insurance company (barring catastrophic events)
- A binomial random variable (sum of binary r.v.'s)
- Gaussian plume model for dispersion of a pollutant: the path a particle in air or water is affected by many small collisions with other particles.

Possible Problems with Normal Models

- Tails of the distribution (e.g., height, wingspan are positive) [Rule of thumb: coefficient of variation less than 0.3]
- Are the additive processes really independent? (e.g., height and sex)
- How many summands are needed to get close to normality?