Lecture Summary

- The normal distribution
- The central limit theorem

Material can be found in Chapter 5.6, 6.3 of DeGroot and Schervish.

The Normal Distribution

Standard normal

$$\mathcal{N}(0,1) : f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

.

Normal with mean μ and variance σ^2

$$\mathcal{N}(0,1) : f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

.

Theorem

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $\alpha X + \beta \sim N(\alpha\mu + \beta, \alpha^2\sigma^2)$

Calculating probabilities with the Normal Distribution

- We want to estimate $P(X \leq a)$ when $X \sim \mathcal{N}(\mu, \sigma^2)$
- No closed form for $\int_{-\infty}^{a} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) dt$
- If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1)$
- $P(X < a) = P(\frac{X-\mu}{\sigma} < \frac{a-\mu}{\sigma}) = \Phi(\frac{a-\mu}{\sigma})$.

Central Limit Theorem

Theorem (Central Limit Theorem)

If the random variables $X_1, \dots, X_n$ form a random sample of size n from a given distribution with mean μ and variance σ^2 $(0 < \sigma^2 < \infty)$, then for each fixed number x

$$lim_{n \to \infty} P(\sqrt{n}\frac{\bar{X}_n - \mu}{\sigma}) = \Phi(x),$$

where Φ denotes the c.d.f. of the standard normal distribution.

Central Limit Theorem

- $S_n = \sum_{i=1}^{n} X_i$, mean $n\mu$, variance $n\sigma^2$.
- $\bar{X}_n = \frac{S_n}{n}$, mean μ variance $\frac{\sigma^2}{n}$.
- $\frac{S_n}{\sqrt{n}}$, mean $\mu\sqrt{n}$, variance σ^2.
- $Zn = \frac{S_n - n\mu}{\sqrt{n}\sigma} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}$, mean 0, variance 1.

Example

- You are doing a poll on "ratio of people agree with the lockdown measures".
- True ratio: p, estimate $\bar{X}_n$
- No guarantee for finding exactly p, so

$$P(|\bar{X}_n - p| \geq 0.01) \leq 0.05$$

- Apply Chebysev inequality with $t = 0.01$: n = 50,000.
- Apply CLT: ?