Chapter 4 Summary

Sampling Distributions

- If $X_1, \ldots, X_n$ is an i.i.d. sample from some distribution and $Y = h(X_1, \ldots, X_n)$ is a random variable determined from the sample by some function h, then the distribution of Y is called a *sampling distribution*.

Convergence of Random Variables

- Three notions of convergence of random variables:
  - $X_n \xrightarrow{a.s.} Y$ or X_n *converges to* Y *almost surely* or *with probability one* as $n \to \infty$ if
  $$\mathsf{P}\left(\lim_{n\to\infty} X_n = Y\right) = 1 \ .$$
  - $X_n \xrightarrow{P} Y$ or X_n *converges to* Y *in probability* as $n \to \infty$ if
  $$\lim_{n\to\infty} \mathsf{P}\left(|X_n - Y| \geq \epsilon\right) = 0$$
  for any $\epsilon > 0$.
  - $X_n \xrightarrow{D} Y$ or X_n *converges to* Y *in distribution* as $n \to \infty$ if
  $$\lim_{n\to\infty} F_{X_n}(x) = F_Y(x)$$
  for all x such that $\mathsf{P}(Y = x) = 0$ where F_Z is the cumulative distribution function of Z.
- These notions of convergence are not identical.
  - $X_n \xrightarrow{a.s.} Y \quad \Rightarrow \quad X_n \xrightarrow{P} Y$ as $n \to \infty$ but the converse is not true in general.
  - $X_n \xrightarrow{P} Y \quad \Rightarrow \quad X_n \xrightarrow{D} Y$ as $n \to \infty$ but the converse is not true in general.
- If $m_Z(s)$ is the moment generating function of random variable Z then
$$\lim_{n\to\infty} m_{X_n}(s) = m_Y(s) \quad \Rightarrow \quad X_n \xrightarrow{D} Y$$
provided the moment generating functions exist.

Laws of Large Numbers

- If $X_1, \ldots, X_n$ is an i.i.d. sample from a distribution with finite mean $\mu = \mathsf{E}(X_i)$, then $M_n = (X_1 + \cdots + X_n)/n$ is the sample mean and $\mathsf{E}(M_n) = \mu$.
- The *weak law of large numbers* says that if we also have $\mathsf{Var}(X_i) = \sigma^2$, then $M_n \xrightarrow{P} \mu$.
- The *strong law of large numbers* says that $M_n \xrightarrow{a.s.} \mu$.

Central Limit Theorem

- If $X_1, \ldots, X_n$ is an i.i.d. sample, $\mathsf{E}(X_i) = \mu$, $\mathsf{Var}(X_i) = \sigma^2 < \infty$, $S_n = \sum_{i=1}^n X_i$, and $M_n = S_n/n$.
- $\mathsf{E}(S_n) = n\mu$, $\mathsf{Var}(S_n) = n\sigma^2$.
- $\mathsf{E}(M_n) = \mu$, $\mathsf{Var}(M_n) = \sigma^2/n$.
- The standardized random variable (mean 0, variance 1) is
$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{M_n - \mu}{\sigma/\sqrt{n}} \ .$$
- If $Z \sim \mathrm{N}(0,1)$, then $Z_n \xrightarrow{D} Z$ as $n \to \infty$.
- For sufficiently large n,
$$\mathsf{P}(S_n \leq c) = \mathsf{P}\left(Z_n \leq \frac{c - n\mu}{\sigma\sqrt{n}}\right) \approx \Phi\left(\frac{c - n\mu}{\sigma\sqrt{n}}\right) \ .$$
- For sufficiently large n,
$$\mathsf{P}(M_n \leq c) = \mathsf{P}\left(Z_n \leq \frac{c - \mu}{\sigma/\sqrt{n}}\right) \approx \Phi\left(\frac{c - \mu}{\sigma/\sqrt{n}}\right) \ .$$

Distributions Based on the Normal

- Sums of independent normal random variables are normal. If $X_i \sim \mathrm{N}(\mu_i, \sigma_i^2)$ for $i = 1, 2, \ldots, n$ and $Y = \sum_{i=1}^n X_i$,
$$Y \sim \mathrm{N}\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right) \ .$$
- **χ^2 distributions:** If $Z_1, \ldots, Z_n \sim$ i.i.d. $\mathrm{N}(0,1)$ and $X = Z_1^2 + \cdots + Z_n^2$, then we say $X \sim \chi^2(n)$.
  - The $\chi^2(n)$ distribution and the Gamma$(n/2, 1/2)$ distributions are identical.
  - $$f(x) = \frac{(1/2)^{n/2}}{\Gamma(n/2)} x^{n/2-1} \mathrm{e}^{-x/2} \text{ for } x > 0$$
  - $\mathsf{E}(X) = n$, $\mathsf{Var}(X) = 2n$.
- **t distributions:** If $Z \sim \mathrm{N}(0,1)$, $Y \sim \chi^2(n)$, and Z and Y are independent, then $X = Z/\sqrt{Y/n} \sim t(n)$.
  - The $t(1)$ distribution is known also as the Cauchy distribution.
  - $$f(x) = \frac{\Gamma((n+1)/2)}{\Gamma(1/2)\Gamma(n/2)\sqrt{n}} \left(1 + \frac{x^2}{n}\right)^{-(n+1)/2} \quad \text{for } x \in \mathbb{R}$$
  - $\mathsf{E}(X) = 0$ for $n > 1$, $\mathsf{Var}(X) = n/(n-2)$ for $n > 2$.
- **F distributions:** If $W \sim \chi^2(m)$, $Y \sim \chi^2(n)$, and W and Y are independent, then $X = \frac{W/m}{Y/n} \sim F(m,n)$.
  - $$f(x) = \frac{\Gamma((m+n)/2)}{\Gamma(m/2)\Gamma(n/2)} \left(\frac{m}{n}\right)^{m+1} x^m \left(1 + \frac{mx}{n}\right)^{-(m+n)/2}$$
  for $x > 0$.
  - $\mathsf{E}(X) = n/(n-2)$ for $n > 2$
  - $\mathsf{Var}(X) = 2n^2(m+n-2)/(m(n-2)^2(n-4))$ for $n > 4$.