Statistics for Data Science - 2

Week 8 Notes

Statistics from samples and Limit theorems

1. **Moment generating function (MGF):**
   Let X be a zero-mean random variable ($E[X] = 0$). The MGF of X, denoted $M_X(\lambda)$, is a function from $\mathbb{R}$ to $\mathbb{R}$ defined as

   $$M_X(\lambda) = E[e^{\lambda X}]$$

   - $$\begin{aligned} M_X(\lambda) &= E[e^{\lambda X}] \\ &= E[1 + \lambda X + \frac{\lambda^2 X^2}{2!} + \frac{\lambda^3 X^3}{3!} + \ldots] \\ &= 1 + \lambda E[X] + \frac{\lambda^2}{2!} E[X^2] + \frac{\lambda^3}{3!} E[X^3] + \ldots \end{aligned}$$

     That is coefficient of $\frac{\lambda^k}{k!}$ in the MGF of X gives the kth moment of X.
   - If $X \sim \text{Normal}(0, \sigma^2)$ then, $M_X(\lambda) = e^{\lambda^2\sigma^2/2}$
   - Let $X_1, X_2, \ldots, X_n \sim$ i.i.d. X and let $S = X_1 + X_2 + \ldots + X_n$, then

     $$M_S(\lambda) = (E[e^{\lambda X}])^n = [M_X(\lambda)]^n$$

     It implies that MGF of sum of independent random variables is product of the individual MGFs.

2. **Central limit theorem:** Let $X_1, X_2, \ldots, X_n \sim$ iid X with $E[X] = \mu, \text{Var}(X) = \sigma^2$. Define $Y = X_1 + X_2 + \ldots + X_n$. Then,

   $$\frac{Y - n\mu}{\sqrt{n}\sigma} \approx \text{Normal}(0, 1).$$

3. **Gamma distribution:**
   $X \sim \text{Gamma}(\alpha, \beta)$ if PDF $f_x(x) \propto x^{\alpha-1}e^{-\beta x}, \quad x > 0$

   - $\alpha > 0$ is a shape parameter.
   - $\beta > 0$ is a rate parameter.
   - $\theta = \frac{1}{\beta}$ is a scale parameter.
   - Mean, $E[X] = \frac{\alpha}{\beta}$
   - Variance, $\text{Var}(X) = \frac{\alpha}{\beta^2}$

4. **Beta distribution:**
$X \sim \text{Beta}(\alpha, \beta)$ if PDF $f_x(x) \propto x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1$

- $\alpha > 0, \beta > 0$ are the shape parameters.
- Mean, $E[X] = \dfrac{\alpha}{\alpha+\beta}$
- Variance, $\text{Var}(X) = \dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

5. **Cauchy distribution:**
$X \sim \text{Cauchy}(\theta, \alpha^2)$ if PDF $f_x(x) \propto \dfrac{1}{\pi}\dfrac{\alpha}{\alpha^2+(x-\theta)^2}$

- θ is a location parameter.
- $\alpha > 0$ is a scale parameter.
- Mean and variance are undefined.

6. **Some important results:**

- Let $X_i \sim \text{Normal}(\mu_i, \sigma_i^2)$ are independent and let $Y = a_1X_1 + a_2X_2 + \ldots a_nX_n$, then
$$Y \sim \text{Normal}(\mu, \sigma^2)$$
where $\mu = a_1\mu_1 + a_2\mu_2 + \ldots a_n\mu_n$ and $\sigma^2 = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \ldots a_n^2\sigma_n^2$
That is linear combinations of i.i.d. normal distributions is again a normal distribution.

- Sum of n i.i.d. $\text{Exp}(\beta)$ is $\text{Gamma}(n, \beta)$.

- Square of $\text{Normal}(0, \sigma^2)$ is $\text{Gamma}\left(\dfrac{1}{2}, \dfrac{1}{2\sigma^2}\right)$.

- Suppose $X, Y \sim$ i.i.d. $\text{Normal}(0, \sigma^2)$. Then, $\dfrac{X}{Y} \sim \text{Cauchy}(0, 1)$.

- Suppose $X \sim \text{Gamma}(\alpha, k), Y \sim \text{Gamma}(\beta, k)$ are independent random variables, then $\dfrac{X}{X+Y} \sim \text{Beta}(\alpha, \beta)$.

- Sum of n independent $\text{Gamma}(\alpha, \beta)$ is $\text{Gamma}(n\alpha, \beta)$.

- If $X_1, X_2, \ldots, X_n \sim$ i.i.d. $\text{Normal}(0, \sigma^2)$, then $X_1^2 + X_2^2 + \ldots + X_n^2 \sim \text{Gamma}\left(\dfrac{n}{2}, \dfrac{1}{2\sigma^2}\right)$.

- Gamma$\left(\frac{n}{2}, \frac{1}{2}\right)$ is called Chi-square distribution with n degrees of freedom, denoted χ^2_n.

- Suppose $X_1, X_2, \ldots, X_n \sim$ i.i.d. Normal(μ, σ^2). Suppose that $\overline{X}$ and S^2 denote the sample mean and sample variance, respectively, then
  (i) $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$
  (ii) $\overline{X}$ and S^2 are independent.