Course 8

Normal (Gaussian)Distribution正态分布(高斯分布)

- If $E(X) = \mu$ and $V(X) = \sigma^2$, $X \sim N(\mu, \sigma^2)$
- **pdf**:
  $f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (-\infty < x < \infty)$

standard normal distribution标准正态分布

The normal distribution with parameter values $\mu = 0$ and $\sigma = 1$ is called the **standard normal distribution**.A random variable having a standard normal distribution is called a **standard normal random variable** and will be denoted by Z.

- **pdf**:
  $f(z;\mu,\sigma) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, Z \sim N(0,1)$
- **cdf**:
  $\Phi(z) = \int_{-\infty}^{z} f(t)\mathrm{d}t = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\mathrm{d}t$
  $\Phi(-z) = 1 - \Phi(z)$
  $\Phi(0) = 0.5$
  $P(|x| \le z) = 2\Phi(z) - 1$
  $P(|x| \ge z) = 2[1 - \Phi(z)]$

  > Proof of $p(|x| \le z) = 2\Phi(z) - 1$:
  >
  > $P(-z \le x \le z) = \Phi(z) - \Phi(-z) = \Phi(z) - (1 - \Phi(z)) = 2\Phi(z) - 1$

  $p = F(\eta) - F(z)$

Z_α notation Z_α表示法

z_α will denote the values on the measurement axis for which α of the area under the z curve lies to the right of z_α.

Z_α is the $100(1-\alpha)$th percentile of the standard normal distribution

- $Z = \frac{X - \mu}{\sigma}$
- $X = \mu + Z\sigma$

> **Proof**:
>
> $X \sim N(\mu, \sigma^2)$
> $F(X) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x$
> Let $\frac{x-\mu}{\sigma} = t$, then $x = \mu + \sigma t, \mathrm{d}x = \sigma\mathrm{d}t$
> $F(X) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-\frac{t^2}{2}}\sigma\mathrm{d}t$

STANDARDIZATION标准化

利用$P(a \le X \le b)$构造$P(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma})$，接下来就可以查**Standard normal density curve**的表了

- $P(a \le X \le b) = P(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}) = \Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})$
- $P(X \le a) = \Phi(\frac{a-\mu}{\sigma})$
- $P(X \ge b) = 1 - \Phi(\frac{b-\mu}{\sigma})$

The Normal Approximation to the Binomial Distribution利用正态分布逼近二项分布

- **Rule**: In practice, the approximation is adequate provided that both $np \geq 10$ and $nq \geq 10$.(where $q = 1 - p$)
- $p(X \leq x) = B(x; n, p) \approx \text{(area under the normal curve to the left of } x + 0.5) = \Phi\left(\frac{x + 0.5 - np}{\sqrt{npq}}\right)$
- **Caution**: the $a - 1$

  $p(a \leq X \leq b) = p(a - 1 < X \leq b) = B(b; n, p) - B(a - 1; n, p) \approx p(\frac{(a-1) + 0.5 - np}{\sqrt{npq}} \leq Z \leq \frac{b + 0.5 - np}{\sqrt{npq}}) = \Phi\left(\frac{b + 0.5 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{(a-1) + 0.5 - np}{\sqrt{npq}}\right)$

Homework

Section 4.3 30, 44, 48, 53, 56