MATH 2411 T1B Tutorial 5

HOU Zhen

zhouah@connect.ust.hk

March 17, 2026

Binomial Distribution $X \sim B(n, p)$

- $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, 1, \ldots, n$.
- For a $B(n, p)$ rv X, let $Z_1, Z_2, \ldots, Z_n$ be the Bernoulli rv from the result of each trial, then $X = Z_1 + Z_2 + \cdots + Z_n$. Importantly, Z_i are independent and have the same distribution $B(1, p)$.
- $\mathbb{E}[X] = np$, $\mathrm{Var}[X] = np(1-p)$.
- Let $X_1 \sim B(n_1, p)$ and $X_2 \sim B(n_2, p)$, if X_1 and X_2 are independent, then $X_1 + X_2 \sim B(n_1 + n_2, p)$.

Exercise

Suppose we play a game where we start with c dollars. On each play of the game you either double or halve your money, with equal probability. What is your expected fortune after n plays?

Hint: $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$.

Poisson Distribution $X \sim P(\lambda)$

- $P(X = k) = \dfrac{e^{-\lambda}\lambda^k}{k!}$, $k = 0, 1, 2, \ldots$.
- $\mathbb{E}[X] = \lambda$, $\mathrm{Var}[X] = \lambda$.
- $n \to \infty, p \to 0, np \to \lambda$, then $B(n, p) \to P(\lambda)$.
- Let $X_1 \sim P(\lambda_1)$ and $X_2 \sim P(\lambda_2)$, if X_1 and X_2 are independent, then $X_1 + X_2 \sim P(\lambda_1 + \lambda_2)$.

Exercise

Prove that let $X_1 \sim P(\lambda_1)$ and $X_2 \sim P(\lambda_2)$, if X_1 and X_2 are independent, then $X_1 + X_2 \sim P(\lambda_1 + \lambda_2)$.

R Example

In R, we can use `dpois`, `ppois`, `qpois`, `rpois` to calculate the pmf, cdf, quantile function, and generate random numbers of Poisson distribution.

```
dpois(0:3, lambda=2)
ppois(0:3, lambda=2)
qpois(0.5, lambda=2)
rpois(10, lambda=2)
```

Normal Distribution

- $X \sim N(\mu, \sigma^2)$, $f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$.
- $\mathbb{E}[X] = \mu$, $\mathrm{Var}[X] = \sigma^2$.
- $X \sim N(0,1)$ is called standard normal distribution.
- $X \sim N(\mu, \sigma^2)$, then $aX + b \sim N(a\mu + b, a^2\sigma^2)$.
- $X \sim N(\mu, \sigma^2)$, then $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$.
- $X_i \sim N(\mu_i, \sigma_i^2), i = 1, 2,$ and X_1, X_2 are independent, then $X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.
- $\Phi(x)$ often denotes the cdf of standard normal distribution. $\Phi(x)$ is strictly increasing and continuous so that $\Phi^{-1}(x)$ exists. The q-th quantile of standard normal distribution can be expressed by $\Phi^{-1}(q)$.

Exercise

$X \sim N(0, \sigma^2)$. Find the cdf of $Y = X^2$.

R Example

In R, we can use `dnorm`, `pnorm`, `qnorm`, `rnorm` to calculate the pdf, cdf, quantile function, and generate random numbers of normal distribution.

```
dnorm(0:3, mean=2, sd=1)
pnorm(0:3, mean=2, sd=1)
qnorm(0.5, mean=2, sd=1)
rnorm(10, mean=2, sd=1)
```

Estimation

Key terms in Statistics

- **UNKNOWN Population**: An unknown distribution of the r.v. X.
- **Sample**: A collection of data of X.
- **Parameter**: e.g. — μ_X and σ_X^2.
- **Statistic**: e.g. — $\bar{x}$, s_{n-1}^2 and s_n^2.

Basic Concepts

- Population distribution $X \sim F_\theta(x)$ with parameter(s) θ.
- Sample $\mathcal{D} = \{X_1, \ldots, X_n\}$. In our course, we often assume i.i.d. sample. For a particular data, the samples are values/realizations of these rvs, denoted by $x_1, \ldots, x_n$.
- Parameter θ: a number that describes the population.
- Estimator $\hat{\theta}(\mathcal{D})$: a rule using the sample to estimate the parameter.

Evaluating an Estimator

- **Accuracy: Bias**: $\text{Bias}(\hat{\theta}, \theta) = \mathbb{E}[\hat{\theta}] - \theta$.
- **Precision: Variance**: $\text{Var}(\hat{\theta}) = \mathbb{E}[(\hat{\theta} - \mathbb{E}[\hat{\theta}])^2]$.
- **Mean Squared Error**:
  $\text{MSE}(\hat{\theta}) = \mathbb{E}[(\hat{\theta} - \theta)^2] = \text{Var}(\hat{\theta}) + (\text{Bias}(\hat{\theta}, \theta))^2$.

Exercise

Let $X_1, \ldots, X_n$ be samples from uniform distribution $U(0, \theta)$. Find c such that $cX_{(n)}$ is an unbiased estimator of θ. Here $X_{(n)} = \max\limits_{i=1,\ldots,n} X_i$.