Department of Mathematics
College of William & Mary

MATH 451/551

Chapter 4. Common Discrete Distributions

4.1 Bernoulli Distribution

GuanNan Wang
gwang01@wm.edu

Examples & Notations

Examples

- **Bernoulli distribution** arises in clinical trials, athletics, polling
- **Poisson distribution** arises in radioactive decay, arrival processes, quality control
- **Normal distribution** arises in agricultural yields, IQ scores, heights of children

Notations

- the support is $\mathcal{A}$
- the parameter space is Ω
- the probability mass function is $f(x) = P(X = x)$
- the cumulative distribution function is $F(x) = P(X \leq x)$
- the moment generating function is $M(t) = E\left(e^{tX}\right)$
- the population mean is $E(X) = \mu$
- the population variance is $V(X) = \sigma^2$

Bernoulli Distribution

- The **Bernoulli distribution** models a **binary** random variable X with support $\mathcal{A} = \{x|x = 0, 1\}$.

success : $X=1$
failure : $X=0$

- A discrete random variable X with probability mass function

$$f(x) = \begin{cases} 1-p & x = 0 \\ p & x = 1 \end{cases}$$

for $0 < p < 1$ is a Bernoulli (p), $Ber(p)$ random variable.
- **Parameter Space:** $\Omega = \{p|0 < p < 1\}$.
- **Compact way to write** $f(x)$**:** $f(x) = p^x(1-p)^{1-x}, \quad x = 0, 1.$
- **Bernoulli Trial:** A single observation of a Bernoulli random variable.
- **Outcomes of a Bernoulli trial are generically called:**
  - success when $X = 1$
  - failure when $X = 0$

Mean

$f(x) = p^{x}(1-p)^{1-x}$, $\mathcal{X} = \{0,1\}$

$E(X) = \sum_{\mathcal{X}} x f(x) = \underline{0 * f(0)} + 1 * f(1) = 1 * p^{1}(1-p)^{1-1} = p$

Variance

$$V(X) = E(X^2) - \{E(X)\}^2$$

$$E(\underbrace{X^2}_{g(X)}) = \sum_x x^2 f(x) = 0^2 * f(0) + 1^2 * f(1) = p$$

$$V(X) = E(X^2) - \{E(X)\}^2 = p - p^2 = p(1-p)$$

MGF

$$M(t) = E(\underbrace{e^{tX}}_{g(X)}) = \sum_x e^{tx} f(x) = e^{t\cdot 0} f(0) + e^{t\cdot 1} f(1)$$

$$= (1-p) + e^t \cdot p = 1-p+e^t\cdot p \,, \quad -\infty < t < \infty$$

$$E(X) = \frac{d\,M(t)}{dt}\bigg|_{t=0} = p\cdot e^t \Big|_{t=0} = p$$

$$E(X^2) = \frac{d^2 M(t)}{dt^2}\bigg|_{t=0} = p\cdot e^t \Big|_{t=0} = p$$

$$V(X) = E(X^2) - \{E(X)\}^2 = p(1-p)$$

Skewness

$$\frac{X-\mu}{\sigma} = \frac{X-p}{\sqrt{p(1-p)}}$$

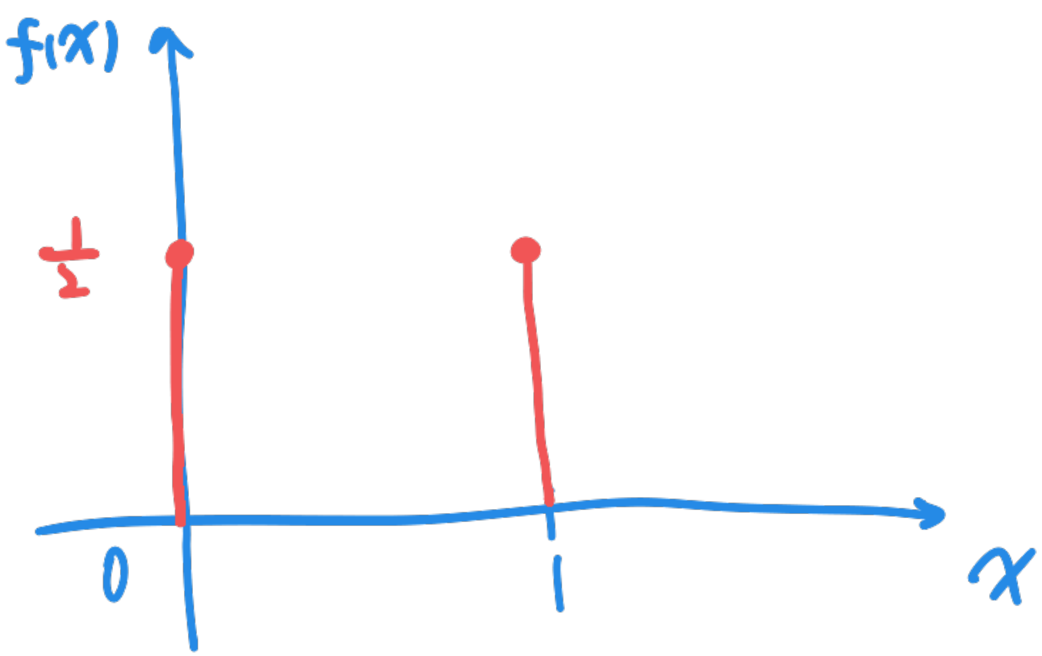
$$E\left\{\left(\frac{X-\mu}{\sigma}\right)^3\right\} = E\left\{\left(\frac{X-p}{\sqrt{p(1-p)}}\right)^3\right\} = \frac{1-2p}{\sqrt{p(1-p)}}$$

$$E\left\{\left(\frac{X-\mu}{\sigma}\right)^4\right\} = E\left\{\left(\frac{X-p}{\sqrt{p(1-p)}}\right)^4\right\} = -3 + \frac{1}{p(1-p)}$$

Remarks

follow

- Shorthand: $X \sim Bernoulli(p)$ or $X \sim Ber(p)$.
- Important special case when $p = \frac{1}{2}$ (fair coin toss)
  - population mean is $\mu = \frac{1}{2}$
  - population variance is $\sigma^2 = \frac{1}{4}$ (largest possible) $\sigma^2 = p(1-p)$
  - population skewness is 0 (symmetric)
  - population kurtosis is 1 when $p = \frac{1}{2}$ (the smallest population kurtosis possible for any random variable)
- Repeated Bernoulli trials form the basis for next three distributions
  - binomial
  - geometric
  - negative binomial



THANK YOU!