EE2012/ST2334 Discussion 4*

Liu Zichen

March 29, 2019

Discrete distributions: Discrete uniform distribution, Bernoulli and Binomial distribution, Negative binomial distribution, Poisson distribution (and its approximation to Binomial distribution).

Continuous distributions: Continuous uniform distribution, Exponential distribution, Normal distribution (and its approximation to Binomial distribution).

1. **[Discrete uniform]** Equal probability for all discrete values. $f_X(x) = 1/k, \quad x = x_1, x_2, \cdots, x_k$, and 0 otherwise. Its mean and variance:

$$\mu = E(X) = \sum x f_X(x) = \sum_{i=1}^{k} x_i \frac{1}{k} = \frac{1}{k} \sum_{i=1}^{k} x_i \tag{1}$$

$$\sigma^2 = V(X) = \sum (x-\mu)^2 f_X(x) = \frac{1}{k} \sum_{i=1}^{k} (x_i - \mu)^2 \quad (= E\left(X^2\right) - \mu^2 = \frac{1}{k} \left(\sum_{i=1}^{k} x_i^2 \right) - \mu^2) \tag{2}$$

2. **[Bernoulli and Binomial]** Random experiments with only **two possible outcomes** are defined as Bernoulli experiments. $f_X(x) = p^x (1-p)^{1-x}, \quad x = 0, 1$, where $0 < p < 1$. We can also denoe as $X \sim Ber(p)$.

$$\mu = E(X) = p \tag{3}$$

$$\sigma^2 = V(X) = p(1-p) \tag{4}$$

If we take the Bernoulli trials for **n** times, with each trial being *independent*, and observe **x** times of success. We say the random variable X, where x is take from, is defined to have a binomial distribution: $\Pr(X = x) = f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$, for $x = 0, 1, \cdots, n$ and $0 < p < 1$. Also denote as $X \sim B(n, p)$. Notice that when $n = 1$, it becomes Bernoulli distribution.

$$\mu = E(X) = np \tag{5}$$

$$\sigma^2 = V(X) = np(1-p) \tag{6}$$

3. **[Negative binomial]** Let X be a random variable that represents the number of trials to produce k successes in a sequence of independent Bernoulli trials. X is said to follow a Negative Binomial distribution, namely $X \sim NB(k, p)$: $\Pr(X = x) = f_X(x) = \binom{x-1}{k-1} p^k q^{x-k}$ for $x = k, k+1, k+2, \cdots$.

$$E(X) = \frac{k}{p} \tag{7}$$

$$\text{Var}(X) = \frac{(1-p)k}{p^2} \tag{8}$$

Notice that the number of trials that are required to have the *first* success is known to follow a special case of negative binomial distribution called *geometric distribution.*

4. **[Poisson]** Experiments yielding numerical values of a random variable X, *the number of successes occurring during a given time interval or in a specified region*, are called Poisson experiments. And the number of successes X in a Poisson experiment is called a Poisson random variable, $X \sim Poisson(\lambda)$: $f_X(x) = \Pr(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$, for $x = 0, 1, 2, 3, \cdots$.

$$E(X) = \lambda \tag{9}$$

$$V(X) = \lambda \tag{10}$$

Recall the Binomial distribution defined in (2), suppose that $n \to \infty$ and $p \to 0$ in such a way that $\lambda = np$ remains constant. We have X being approximated by a Poisson distribution:

$$\lim_{\substack{p \to 0 \\ n \to \infty}} \Pr(X = x) = \frac{e^{-np}(np)^x}{x!}$$

If $p \to 1$, we can still use the approximation by interchanging the definition of success and failure.

*Some examples from bookmarked pages of the lecture notes

5. **[Continuous uniform]** A continuous random variable, which is uniformly distributed over the interval $[a, b]$, $-\infty < a < b < \infty$. $f_X(x) = \frac{1}{b-a}$, for $a \leq x \leq b$, and 0 otherwise.

$$E(X) = \frac{a+b}{2} \tag{11}$$

$$V(X) = \frac{1}{12}(b-a)^2 \tag{12}$$

6. **[Exponential]** A continuous random variable X assuming all nonnegative values is said to have an exponential distribution with parameter $\alpha > 0$ if its probability density function is given by $f_X(x) = \alpha e^{-\alpha x}$, for $x > 0$. Denote as $X \sim \text{Exp}(\alpha)$

$$E(X) = \frac{1}{\alpha} \tag{13}$$

$$V(X) = \frac{1}{\alpha^2} \tag{14}$$

No Memory Property of Exponential Distribution: for any two positive numbers s and t, $\Pr(X > s + t | X > s) = \Pr(X > t)$. *Meaning*: If X denotes the life length of a bulb, given that the bulb has lasted s time units, then the probability of it lasting for the next t time units is the same as the probability that it would last for the first t time units as brand new.

Another note is that the exponential distribution is frequently used as a model for the *distribution of times between the occurrence of successive events* such as customers arriving at a service facility or calls coming in to a switchboard.

7. **[Gaussian]** The PDF of Gaussian (normal) distribution is: $f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$, $-\infty < x < \infty$, where $-\infty < \mu < \infty$ and $\sigma > 0$. Denote as $X \sim N\left(\mu, \sigma^2\right)$.

$$E(X) = \mu \tag{15}$$

$$V(X) = \sigma^2 \tag{16}$$

To obtain the standardized Gaussian, let $Z = \frac{(X-\mu)}{\sigma}$, and result in $Z \sim N(0,1)$, $f_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{Z^2}{2}\right)$.

Statistical table: gives the values $\Phi(z)$ of a given z, where $\Phi(z)$ is the cumulative distribution function of a standardized Normal random variable Z.

$$\begin{gathered} \Phi(z) = \Pr(Z \leq z) \\ 1 - \Phi(z) = \Pr(Z > z) \end{gathered} \tag{17}$$

Recall the Binomial distribution defined in (2), suppose that $n \to \infty$ and $p \to \frac{1}{2}$ (or even when n is small and p is not extremely close to 0 or 1), we have X being approximated by a Gaussian distribution with mean $\mu = np$ and variance $\sigma^2 = np(1-p)$:

$$Z = \frac{X - np}{\sqrt{npq}} \text{ is approximately } \sim N(0,1)$$