VE401 Probabilistic Methods in Eng.
RC 3

CHEN Xiwen

UM-SJTU Joint Institute

March 20, 2020

Table of contents

Random Variables and Probability Density Function

Definition. Let S be a sample space and Ω a <u>countable</u> subset of $\mathbb{R}$. A ***discrete random variable*** is a map

$$X : S \to \Omega$$

together with a function

$$f_X : \Omega \to \mathbb{R}$$

having the properties that

(i) $f_X(x) \geq 0$ for all $x \in \Omega$ and

(ii) $\sum_{x \in \Omega} f_X(x) = 1$.

The function f_X is called the ***probability density function*** or ***probability distribution*** of X. A random variable is given by the pair (X, f_X).

Cumulative Distribution Function

Definition. The ***cumulative distribution function*** of a random variable is defined as

$$F_X : \mathbb{R} \to \mathbb{R}, \qquad F_X(x) := P[X \leq x].$$

For a discrete random variable,

$$F_X(x) = \sum_{y \leq x} f_X(y).$$

Expectation and Variance

Definition. Let (X, f_X) be a discrete random variable.

- The ***expected value*** or ***expectation*** of X is

$$\mu_X = \mathrm{E}[X] := \sum_{x \in \Omega} x \cdot f_X(x),$$

provided that the sum (possibly series, if Ω is infinite) on the right converges absolutely.

- The ***variance*** is defined by

$$\sigma_X^2 = \mathrm{Var}[X] := \mathrm{E}\left[(X - \mathrm{E}[X])^2\right]$$

which is defined as long as the right-hand side exists.

- The ***standard deviation*** is $\sigma_X = \sqrt{\mathrm{Var}[X]}$.

Properties

- <u>Expectation</u>.

  (a). Suppose $\varphi : \Omega \to \mathbb{R}$ is some function, then

  $$\mathrm{E}[\varphi \circ X] = \sum_{x \in \Omega} \varphi(x) \cdot f_X(x).$$

  (b). $\mathrm{E}[aX + bY + c] = a\mathrm{E}[X] + b\mathrm{E}[Y] + c$, where $a, b, c \in \mathbb{R}$ and X, Y are random variables.

  (c). $\mathrm{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathrm{E}[X_i]$, if each expectation exists.

  (d). If $X_1, \dots, X_n$ are independent random variables with finite expectations, and $g_i, i = 1, \dots, n$ are functions, then

  $$\mathrm{E}\left[\prod_{i=1}^{n} X_i\right] = \prod_{i=1}^{n} \mathrm{E}[X_i], \quad \mathrm{E}\left[\prod_{i=1}^{n} g_i(X_i)\right] = \prod_{i=1}^{n} \mathrm{E}[g_i(X_i)].$$

Properties

- Variance.
  (a). $\mathrm{Var}[X] = \mathrm{E}[X^2] - \mathrm{E}[X]^2$.
  (b). $\mathrm{Var}[aX + b] = a^2\mathrm{Var}[X]$, where $a, b \in \mathbb{R}$.
  (c). If $X_1, \ldots, X_n$ are independent random variables, then

$$\mathrm{Var}\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2 \mathrm{Var}[X_i].$$

Note. If X and Y are not independent, then according to definitions,

$$\begin{aligned}\mathrm{Var}[X+Y] &= \mathrm{E}\left[(X + Y - (\mu_X + \mu_Y))^2\right] \\ &= \mathrm{E}\left[(X - \mu_X)^2\right] + \mathrm{E}\left[(Y - \mu_Y)^2\right] + \\ &\qquad + 2\mathrm{E}\left[(X - \mu_X)(Y - \mu_Y)\right] \\ &\neq \mathrm{Var}[X] + \mathrm{Var}[Y].\end{aligned}$$

Ordinary and Central Moments

Definition. The n^{th} ***(ordinary) moments*** of a random variable X is given by

$$\mathrm{E}[X^n], \qquad n \in \mathbb{N}.$$

The n^{th} ***central moments*** of X is given by

$$\mathrm{E}\left[\left(\frac{X-\mu}{\sigma}\right)^n\right], \qquad \text{where } n = 3, 4, 5, \ldots$$

Moment-Generating Function

Definition. Let (X, f_X) be a random variable and such that the sequence of moments $\mathrm{E}[X^n], n \in \mathbb{N}$, exists. If the power series

$$m_X(t) := \sum_{k=0}^{\infty} \frac{\mathrm{E}[X^k]}{k!} t^k$$

has radius of convergence $\varepsilon > 0$, the thereby defined function

$$m_X(t) : (-\varepsilon, \varepsilon) \to \mathbb{R}$$

is called the ***moment-generating function*** for X.

Moment-Generating Function

Theorem. Let $\varepsilon > 0$ be given such that $\mathrm{E}[e^{tX}]$ exists and has a power series expansion in t that converges for $|t| < \varepsilon$. Then the moment-generating function exists and

$$m_X(t) = \mathrm{E}[e^{tX}] \qquad \text{for } |t| < \varepsilon.$$

Furthermore,

$$E[X^k] = \left.\frac{\mathrm{d}^k m_X(t)}{\mathrm{d}t^k}\right|_{t=0}.$$

We can hence calculate the moments of X by differentiating the moment-generating function.

Moment-Generating Function

Properties.

- X is a random variable and $Y = aX + b, a, b \in \mathbb{R}$, then for every t such that $m_X(at)$ is finite,

$$m_Y(t) = e^{bt} m_X(at).$$

- Suppose $X_1, \dots, X_n$ are n independent random variables, then for every value that $m_{X_i}(t)$ is finite for all $i = 1, \dots, n$,

$$m_X(t) = \prod_{i=1}^{n} m_{X_i}(t), \qquad X = X_1 + \cdots + X_n.$$

Moment-Generating Function

Example 1. Suppose that X is a random variable with the moment-generating function

$$m_X : \mathbb{R} \to \mathbb{R}, \qquad m_X(t) = e^{t^2+3t}.$$

Find the mean and variance of X.

Moment-Generating Function

Example 1. Suppose that X is a random variable with the moment-generating function

$$m_X : \mathbb{R} \to \mathbb{R}, \qquad m_X(t) = e^{t^2+3t}.$$

Find the mean and variance of X.

Solution. We calculate

$$m_X'(t) = (2t+3)e^{t^2+3t}, \quad m_X''(t) = (2t+3)^2 e^{t^2+3t} + 2e^{t^2+3t}.$$

Therefore,

$$\mu = m_X'(0) = 3, \quad \sigma^2 = \mathrm{E}[X^2] - \mathrm{E}[X]^2 = m_X''(0) - \mu^2 = 2.$$

Bernoulli Distribution

Definition. A random variable (X, f_X) has a ***Bernoulli distribution*** with parameter $p, 0 < p < 1$ if the probability density function is defined by

$$f_X : \{0,1\} \to \mathbb{R}, \qquad f_X(x) = \begin{cases} 1-p, & \text{if } x = 0, \\ p, & \text{if } x = 1. \end{cases}$$

Interpretation. Describe the probability of success $f_X(1)$ or failure $f_X(0)$ of a trial, given the probability of success is p.

Bernoulli Distribution

Mean, variance, and M.G.F.

- Mean.

$$E[X] = 0 \cdot (1 - p) + 1 \cdot p = p.$$

- Variance.

$$\mathrm{Var}[X] = E[X^2] - E[X]^2 = p - p^2 = p(1 - p).$$

- M.G.F.

$$m_X : \mathbb{R} \to \mathbb{R}, \qquad m_X(t) = (1 - p) + e^t p.$$

Binomial Distribution

Definition. A random variable (X, f_X) has a ***binomial distribution*** with parameter $n \in \mathbb{N} \setminus \{0\}$ and $p, 0 < p < 1$ if it has probability density function

$$f_X : \{0, \dots, n\} \to \mathbb{R}, \qquad f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}.$$

Interpretation. $f_X(x)$ is the probability of obtaining x successes in n independent and identical Bernoulli trials with parameter p.

Binomial Distribution

Mean, variance and M.G.F.

- Mean.

$$E[X] = \sum_{i=1}^{n} E[X_i] = np.$$

- Variance.

$$\mathrm{Var}[X] = \sum_{i=1}^{n} \mathrm{Var}[X_i] = np(1-p).$$

- M.G.F.

$$m_X : \mathbb{R} \to \mathbb{R}, \qquad m_X(t) = E[e^{tX}] = \prod_{i=1}^{n} E[e^{tX_i}] = (1-p+pe^t)^n.$$

Geometric Distribution

Definition. A random variable (X, f_X) has ***geometric distribution*** with parameter $p, 0 < p < 1$ if the probability density function is given by

$$f_X : \mathbb{N} \setminus \{0\} \to \mathbb{R}, \qquad f_X(x) = (1-p)^{x-1}p.$$

Interpretation. $f_X(x)$ is the probability of x failures before the first success in the Bernoulli trials, given the probability of success for each trial is p.

Geometric Distribution

Mean, variance and M.G.F.

- Mean.

$$\mathrm{E}[X] = \frac{1}{p}.$$

- Variance.

$$\mathrm{Var}[X] = \frac{1-p}{p^2}.$$

- M.G.F.

$$m_X : (-\infty, -\ln(1-p)) \to \mathbb{R}, \qquad m_X(t) = \frac{pe^t}{1-(1-p)e^t}.$$

Pascal Distribution

Definition. A random variable (X, f_X) has the ***Pascal distribution*** with parameters $p, 0 < p < 1$ and $r \in \mathbb{N} \setminus \{0\}$ if the probability density function is given by

$$f_X : \{r, r+1, \ldots\} \to \mathbb{R}, \qquad f_X(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}.$$

Interpretation. $f_X(x)$ is the probability of obtaining the r-th success in the x-th Bernoulli trail, given the probability of success for each trail is p.

Pascal Distribution

Mean, variance and M.G.F.

- Mean.

$$\mathrm{E}[X] = \frac{r}{p}.$$

- Variance. Let $q = 1 - p$,

$$\mathrm{Var}[X] = \frac{rq}{p^2}.$$

- M.G.F.

$$m_X : (-\infty, -\ln q) \to \mathbb{R}, \qquad m_X(t) = \frac{(pe^t)^r}{(1 - qe^t)^r}.$$

Negative Binomial Distribution

Definition. A random variable (X, f_X) has the ***negative binomial distribution*** with parameters r and p if the probability density function is given by

$$f_X : \mathbb{N} \to \mathbb{R}, \qquad f_X(x) = \binom{x+r-1}{r-1} p^r (1-p)^x.$$

Interpretation. $f_X(x)$ is the probability of x failures before first obtaining r successes in Bernoulli trials, given the probability for each success is p.

Negative Binomial Distribution

Mean, variance and M.G.F.

- Mean. Let $q = 1 - p$,

$$\mathrm{E}[X] = \frac{rp}{q}.$$

- Variance.

$$\mathrm{Var}[X] = \frac{rp}{q^2}.$$

- M.G.F.

$$m_X : (-\infty, -\ln q) \to \mathbb{R}, \qquad m_X(t) = \frac{p^r}{(1 - qe^t)^r}.$$

Poisson Distribution

Definition. A random variable X has the ***Poisson distribution*** with parameter $k > 0$ if probability density function is given by

$$f_X : \mathbb{N} \to \mathbb{R}, \qquad f_X(x) = \frac{k^x e^{-k}}{x!}$$

Interpretation. $f_X(x)$ is the probability of x arrivals in the time interval $[0, t]$ with arrival rate $\lambda > 0$, and $k = \lambda t$.

"...which describes the occurrence of events that occur at a <u>constant rate</u> and <u>continuous environment</u>."

Poisson Distribution

Interpretation. *Constant rate* and *continuous environment*?

- Continuous environment. Not limited to time intervals, but also subregions of two- or three-dimensional regions or sublengths of a linear distance, and any regions that can be divided into arbitrarily small pieces.
- Constant rate. The probability of an occurrence during each very short interval (region) must be approximately proportional to the length (area, volume) of that interval (region).

Poisson Distribution

Interpretation. *Constant rate* and *continuous environment*?

- Continuous environment. Not limited to time intervals, but also subregions of two- or three-dimensional regions or sublengths of a linear distance, and any regions that can be divided into arbitrarily small pieces.
- Constant rate. The probability of an occurrence during each very short interval (region) must be approximately proportional to the length (area, volume) of that interval (region).

Examples. Poisson process can be used to model

(a). the number of particles that strike a certain target at a constant rate in a particular period;

(b). the number of oocysts that occur in a water supply system given constant rate of occurrence per liter;

and many more.

Poisson Distribution

Mean, variance and M.G.F.

- Mean.

$$\mathrm{E}[X] = k.$$

- Variance.

$$\mathrm{Var}[X] = k.$$

- M.G.F.

$$m_X : \mathbb{R} \to \mathbb{R}, \qquad m_X(t) = e^{k(e^t - 1)}.$$

Distributions Based on Bernoulli Trials

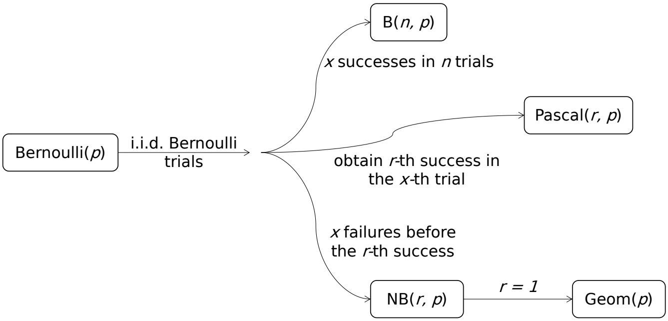

Figure: Connections of distributions based on Bernoulli trials.

Connections of Distributions

- Bernoulli → Binomial. $X_1, \ldots, X_n$ are independent random variables,

$$X_i \sim \text{Bernoulli}(p) \quad \Rightarrow \quad X = X_1 + \cdots + X_n \sim \text{B}(n, p).$$

- Binomial → Binomial. $X_1, \ldots, X_k$ are independent random variables,

$$X_i \sim \text{B}(n_i, p) \quad \Rightarrow \quad X = X_1 + \cdots + X_k \sim \text{B}(n, p),$$

  where $n = n_1 + \cdots + n_k$.

- Geometric → Negative binomial. $X_1, \ldots, X_r$ are independent random variables,

$$X_i \sim \text{Geom}(p) \quad \Rightarrow \quad X = X_1 + \cdots + X_r \sim \text{NB}(r, p).$$

Connections of Distributions

- Negative binomial $\to$ Negative binomial. $X_1, \ldots, X_n$ are independent random variables,

$$X_i \sim \mathrm{NB}(r_i, p) \quad \Rightarrow \quad X = X_1 + \cdots + X_n \sim \mathrm{NB}(r, p),$$

  where $r = r_1 + \cdots + r_n$.

- Poisson $\to$ Poisson. $X_1, \ldots, X_n$ are independent random variables,

$$X_i \sim \mathrm{Poisson}(k_i) \quad \Rightarrow \quad X = X_1 + \cdots + X_n \sim \mathrm{Poisson}(k),$$

  where $k = k_1 + \cdots + k_n$.

Closeness of Binomial Distribution and Poisson Distribution

Theorem. For $n \in \mathbb{N} \setminus \{0\}, 0 < p < 1$, suppose $f(x; n, p)$ denotes the probability density function of binomial distribution with parameters n and p, while $f(x; k)$ denotes the probability density function of Poisson distribution with parameter k. Let $\{p_n\}_{n=1}^{\infty}$ be a sequence of numbers between 0 and 1 such that

$$\lim_{n \to \infty} np_n = k,$$

then

$$\lim_{n \to \infty} f(x; n, p_n) = f(x; k), \qquad \text{for all } x = 0, 1, \ldots$$

This means we can approximate the binomial distribution with Poisson distribution when n is large. A proof can be found in `s2.pdf`.

Exercises

Exercise 1. Suppose Keven plays a game where he has probability p to win in each play. When he wins, his fortune is doubled, and when he loses, his fortune is cut in half. If he begins playing with a given fortune $c > 0$, what is the expected value of his fortune after n independent plays?

Exercises

Exercise 2. For $0 < p < 1$ and $n = 2, 3, \ldots$, determine the value of

$$\sum_{x=2}^{n} x(x-1)\binom{n}{x} p^x (1-p)^{n-x}.$$

Exercises

Exercise 3. Suppose that a book with n pages contains on the average λ misprints per page. What is the probability that there will be at least m pages which contain more than k misprints?

Thanks for your attention!