Important Distributions and Densities

Mark Huiskes, LIACS
mark.huiskes@liacs.nl

Introduction

- This class we will consider a number of important distributions and densities.
- Together they cover quite a few of the elementary probabilistic models often used in practice.
- First we discuss a number of discrete distribution functions, then some continuous density functions. In the last part I'll also show how you can compute the distribution/density function that is a function of a different random variable.
- Next week, some of them will serve as examples for computing the the expectation and variance of distributions.

Discrete Uniform Distribution

- Sample space: $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$
- Distribution function: $m(\omega) = \frac{1}{n}$ for all $\omega \in \Omega$
- Picture:
- Example: Throwing a fair die; drawing a ball from an urn etc.
- Often used in “symmetric” problems: no outcome is more probable than another

Binomial Distribution

- Counts the number of successes in a Bernoulli trials process with parameters n and p
- Sample space: $\Omega = \{1, 2, 3, \ldots, n\}$
- Distribution function: $m(\omega) = \left(\begin{array}{c} n \\ \omega \end{array}\right) p^{\omega}(1-p)^{n-\omega}$, or:

$$b(p,n,k) = \left(\begin{array}{c} n \\ k \end{array}\right) p^k q^{n-k}$$

- Decision tree

Geometric Distribution

- Models the trial of first success in a Bernoulli trials process with parameters n and p
- Sample space: $\Omega = \{1, 2, 3, \ldots\}$
- Let T be the number of the trial at which the first success occurs. [Decision tree]. Then

  P(T=1) = p

  P(T=2) = qp

  P(T=3) = q^2 p

  :

  P(T=n) = q^(n-1) p
- Distribution function: $m(\omega) = (1-p)^{\omega-1}p$ or: $P(T=j) = q^{j-1}p$
- Called "geometric" because of its relation to the geometric series: 1 + s + s^2 + s^3 + …= 1 / (1 – s). [Derive]

Geometric Distribution (more)

- Example: Make assignment 8: $P(T > 5|T > 2) = P(T > 3) = q^3 = 1/8$

Show in the assignment that:

- $P(T>k) = q^k(p + qp + q^2p + \dots) = q^k$

- Memory-less property $P(T > r+s|T>r) = P(T>s) = q^s$

Poisson Distribution (introduction)

- Models the number of random occurrences in an interval, [e.g. the number of incoming customers, or telephone calls.]
- Sample space: $\Omega = \{0, 1, 2, 3, \ldots\}$
- Assumptions:
  - the average rate is a constant: λ
  - The number of occurrences in disjoint intervals are independent
- Approximate the situation for an interval of length t using a binomial probability: n intervals with probability of occurrence $p = \frac{\lambda t}{n}$, as that gives the right rate.

Poisson Distribution (continued)

- The Poisson distribution approximates the binomial distribution for large n and small p

- X: Poisson variable with parameter lambda

  X_n: Approximating binomial variable

with $p = \frac{\lambda}{n}$, we have that

$$P(X = k) = \lim_{n \to \infty} P(X_n = k) = \lim_{n \to \infty} \binom{n}{k} p^k (1-p)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}$$

- Distribution function: $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$

Poisson Distribution (better derivation)

$$\lim_{n\to\infty} (1 - \frac{\lambda}{n})^n = e^{-\lambda}$$

with $p = \frac{\lambda}{n}$, we have that

$$\lim_{n\to\infty} P(X = k) = \lim_{n\to\infty} \left(\begin{array}{c} n \\ k \end{array} \right) p^k (1-p)^{n-k} = \lim_{n\to\infty} \frac{n!}{(n-k)!k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$= \lim_{n\to\infty} \left(\frac{n}{n}\right) \left(\frac{n-1}{n}\right) \cdots \left(\frac{n-k+1}{n}\right) \left(\frac{\lambda^k}{k!}\right) \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}$$

$$= \frac{\lambda^k}{k!} e^{-\lambda}$$

Example

- Printing words. Suppose for each word there is a probability of 1/1000 that a spelling mistake is made. Suppose there are 100 words on a page: what is the probability distribution of the number of mistakes on a page (S)

- Binomial

$$P(S = k) = \binom{100}{k} \frac{1}{1000^k} \left(1 - \frac{1}{1000}\right)^{100-k}$$

- Poisson: $\lambda = np = 100 \times \frac{1}{1000} = \frac{1}{10}$

$$P(S = k) = \frac{.1^k}{k!} e^{-.1}$$

- Probability of at least one spelling mistake:

$$P(S \geq 1) = 1 - P(S = 0) = 1 - e^{-.1} = 0.0952$$

Assignment

- Assignment 18: p = 1/500. Chance a bit hits a particular cookie is 1/500.
- R: #raisins in particular cookie, C: #chips in particular cookie

- lam_R = 600 * 1/500; lam_C = 400 * 1/500

- Any bits: lam_B = 1000 * 1/500.

  Also explain alternative way:

  1-P(R=0, C=0)-P(R=1, C=0)-P(R=0, C=1) + independence, also gives 0.5940

The Continuous Uniform Density

- Random variable U whose value represents the outcome of the experiment consisting of choosing a real number at random from the interval [a, b].
- Density:

$$f(\omega) = \begin{cases} 1/(b-a) & \text{if} \quad a \leq \omega \leq b, \\ 0 & \text{if} \quad \text{otherwise} \end{cases}$$

The Exponential Density

- Often used to model times between independent events that happen at a constant average rate

- Density:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if} \quad x \geq 0, \\ 0 & \text{if} \quad \text{otherwise} \end{cases}$$

- Cumulative distribution function:

$$F(x) = P(T \leq x) = \int_0^x \lambda e^{-\lambda t}\, dt = 1 - e^{-\lambda x}$$

- Memoryless property: P(T>r+s|T>r) = P(T>s)

Relationships with other distributions

- The exponential density is the limit case of the geometric distribution with the same setup as for Poisson
- The Poisson distribution with parameter λ can be simulated by counting how many realizations of an exponential variable with parameter λ fit in a unit interval
- [[The exponential density gives the waiting times for the Poisson case. For instance with a Poisson variable with parameter λt we have;

$$P(X = 0) = e^{-\lambda t}$$

so the probability of waiting a certain time goes down exponentially like in the exponential distribution]]

Normal Density

- According to the book the most important density function. We will see why later.
- Sample space: $\Omega = \mathbb{R}$

- Density function with parameters μ and σ

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

μ : center; σ : spread

- Cumulative distribution $F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(s-\mu)^2}{2\sigma^2}}\, ds$

- The normal density with has a normal density with $\mu = 0$ and $\sigma = 1$ is called the standard normal density:

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

Functions of Random Variables

- Start with an example: Assignment 1

- Now for a general (strictly increasing) function ϕ and $Y = \phi(X)$:

$$F_Y(y) = P(Y \leq y) = P(\phi(X) \leq y) = P(X \leq \phi^{-1}(y)) = F_X(\phi^{-1}(y))$$

- Very similar for strictly decreasing.

- The density function of Y can be determined by differentiating the cumulative distribution function (increasing):

$$f_Y(y) = f_X(\phi^{-1}(y)) \frac{d}{dy} \phi^{-1}(y)$$

Example

- Suppose Z has a standard normal density:

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

- Show that $X = \sigma Z + \mu$ has a normal density with parameters μ and σ :

$$\phi(z) = \sigma z + \mu, \text{ so } \phi^{-1}(x) = \frac{x-\mu}{\sigma}$$

So:

$$F_X(x) = F_Z(\frac{x-\mu}{\sigma})$$

$$f_X(x) = f_Z(\frac{x-\mu}{\sigma}) \cdot \frac{1}{\sigma} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x-\mu}{2\sigma^2}}$$

- Similarly: if X has a normal density with parameters mu and sigma, then Z = (X-mu)/sigma is standard normal

Example with Normal Distribution Table

- P(Z <= 1.56)
- P(Z<= -1.56)

Simulation

- Simulate random variable with a strictly increasing cumulative distribution function $F(y)$

- Use that $Y = F^{-1}(U)$ has cumulative distribution $F(y)$ if U is uniformly distributed on [0,1]:

$$P(Y \leq y) = P(F^{-1}(U) \leq y) = P(U \leq F(y)) = F(y)$$

- So we can simulate values from such a random variable with values $F^{-1}(u)$, with u from the uniform distribution