ECE560: Computer Systems Performance Evaluation

Lecture #6 –

Probability Theory (Part II Review)
- Random Variables

Instructor: Dr. Liudong Xing
Spring 2024

Administration Issues (2/7)

- Homework #2
  - Please download problems from course website
  - Due: **February 12, Monday**
- Project proposal (refer to Guidelines)
  - Due: **February 23, Friday**

Review of Lecture#5

- Probability theory: a basic tool for modeling random / uncertain phenomena
  - Sample spaces & events
  - Axioms of probability
  - Field, σ-field, and probability measure
  - Odds for and odds against
  - Conditional probability & law of total probability
  - Pair-wise vs mutually independence

Related reading: Allen's Ch. 2.0~2.4

Topics

- Random Variable (*r.v.*)
  - **Basic concepts**
  - Discrete *r.v.s*
  - Continuous *r.v.s*

Related reading: Allen's Ch. 2.5~2.6, 3.0~3.2

Random Variable

- Informally, a random variable (*r.v.*) X is a real-valued function from some sample space Ω to *R, i.e., X:* $\Omega \rightarrow R$
- A *r.v.* X maps each outcome ω in Ω *to a* real number $X(\omega) \in R$

- *r.v.* is not *random* or *variable*, but a *function*
- The mapping is not random but deterministic

Random Variable Example

- *"tossing a fair coin three times"*
  - Ω = {TTT; TTH; THT; THH; HTT; HTH; HHT; HHH}
  - Let X be the number of heads tossed in 3 times
  - We can map each outcome in Ω to a real number.

Distribution Function

- Definition: The cumulative distribution function (c.d.f.) or more simply the distribution function F of a *r.v.* X is defined for each real number x, by

$$\begin{aligned} F(x) &= P[\{\omega : \omega \in \Omega \text{ AND } X(\omega) \le x\}] \\ &= P\{X \le x\} \end{aligned}$$

- Property:
  - F is a non-decreasing function

    If $x < y$ then $F(x) <= F(y)$
  - For all $x < y$, $P\{x < X \le y\} = F(y) - F(x)$
  - Also, $\lim_{x\to\infty} F(x) = 1$ $\quad \lim_{x\to-\infty} F(x) = 0$

Discrete vs. Continuous Random Variable

- A r.v. that can take on (at most) a countable number of possible values is said to be a ***discrete*** r.v.

  Example: tossing coin example

- A r.v. that can take on a range of real values (uncountable!) is said to be a ***continuous*** r.v.

  Example: the time of arrival of a packet to a switch
  - In general, the time when a particular event occurs

Discrete vs. Continuous Random Variable (Agenda)

✓ Definitions

- Characteristic functions
  - Probability mass function (p.m.f.) for discrete *r.v.*
  - Probability density function (p.d.f.) for continuous *r.v.*
- Important parameters of *r.v.*s
  - Expectation/Mean: E[X], the *k*th moment
  - Variance/standard deviation
  - Squared coefficient of variation (C.O.V.) and C.O.V.
- Important example *r.v.*s
  - *Discrete*: Bernoulli, Binomial, Geometric, Poisson
  - *Continuous*: Uniform, Normal, Exponential

Discrete Random Variable (1)

- Probability mass function (p.m.f.)

The p.m.f. denoted by $P_X(x)$ for a discrete r.v. X is defined by

$$\begin{aligned} p_X(x) &= P\{X = x\} \\ &= P[\{\omega : \omega \in \Omega \mid X(\omega) = x\}] \end{aligned}$$

- $P_X(x)$ is the probability that the r.v. X takes on a value of x
- $0 <= p_X(x) <= 1$
- Since *r.v.* assigns some real value $x \in$ R to each sample point $\omega \in \Omega$, we must have

$$\sum_{x \in R} p_X(x) = 1 \quad \textbf{or} \quad \sum_{x_i \in T} p_X(x_i) = 1$$

where T = $\{x_1, x_2, \ldots\}$ is the image of X

Discrete Random Variable (2)

- Expectation/Mean/Expected value of X

Let X be a discrete r.v. with p.m.f. $p_X(x)$, the mean of X, denoted by μ=E[X], is defined by

$$\mu = E[X] = \sum_{x_i \in T} (x_i \bullet p_X(x_i))$$

- Expectation of a function of a r.v. X

If X is a discrete r.v. that takes on one of the values $x_i \in T$ with respective probability $p_X(x_i)$, then for any real-valued function g:

$$E[g(X)] = \sum_{x_i \in T} (g(x_i) \bullet p_X(x_i))$$

Discrete Random Variable (3)

- The kth moment of X

The kth moment of X is defined by $E[X^k]$, $\{k = 1, 2, 3, \ldots\}$.

$$E[X^k] = \sum_{x_i \in T} (x_i^k \bullet p_X(x_i))$$

Discrete Random Variable (4)

- Variance of X

If X is a r.v. with mean μ, then the variance of X, denoted by Var(X)= δ^2, is defined by Var(X)= $E[(X-\mu)^2]$

– If X is a discrete r.v., then

$$\boldsymbol{Var}[\boldsymbol{X}] = \boldsymbol{E}[(\boldsymbol{X}-\mu)^2] = \sum_{x_i \in T} [(\boldsymbol{x}_i - \mu)^2 \bullet \boldsymbol{p}_X(\boldsymbol{x}_i)]$$

– An alternative formula:

$$Var[X] = E[X^2] - (E[X])^2 = E[X^2] - \mu^2$$

Discrete Random Variable (5)

- Standard deviation of X

The squared root of the variance Var(X) is called the standard deviation of X:

$$\sigma = SD(X) = \sqrt{Var[X]}$$

- Squared coefficient of variation (C.O.V)

$$C_X^2 = \frac{Var[X]}{E[X]^2} = \frac{\sigma^2}{\mu^2}$$

- Coefficient of variation (C.O.V)

$$COV = \sqrt{\frac{Var[X]}{E[X]^2}} = \frac{\sigma}{\mu}$$

Discrete Random Variable (6)

- <u>Hands-on problem</u>:

  "tossing a fair coin 3 times"

  Let *r.v.* X be the number of heads tossed in 3 times

  – Define the p.m.f.? c.d.f.?

  – Find the mean, and the variance of X?

Discrete Random Variable (Agenda)

✓ Definitions

  ✓ A r.v. that can take on (at most) a countable number of possible values is said to be a ***discrete*** r.v.

✓ Characteristic functions

  ✓ Probability mass function (p.m.f.) for discrete *r.v.*

✓ Important parameters of *r.v.*s

  ✓ Expectation/Mean: E[X], the *k*th moment

  ✓ Variance/standard deviation

  ✓ Squared coefficient of variation (C.O.V.) and C.O.V.

- **<u>Important example *r.v.*s</u>**

  – *Discrete*: Bernoulli, Binomial, Geometric, Poisson

Example Discrete *r.v.*s (1)

- Bernoulli r.v.: has a 0 and 1 as its only possible values
  - Originates from the *Bernoulli trial*, which is a random experiment in which there are only 2 possible outcomes: success (1) or failure (0), with respective probability *p* and *q*, where *p*+*q*=1.
  - P.m.f: $P\{X=1\} = p_X(1) = p_1 = p$
    $P\{X=0\} = p_X(0) = p_0 = q = 1-p$
  - C.d.f:
    $$F(x) = \begin{cases} 0 & x<0 \\ q & 0 \le x < 1 \\ p+q=1 & x \ge 1 \end{cases}$$
  - E[X]: p

Example Discrete *r.v.*s (2)

- Binomial r.v.: is a r.v X that counts the number of successes in the *n* independent Bernoulli trials, each trial has success probability of *p* and failure probability of *q* (*p*+*q*=*1*).
  - P.m.f:
    $P\{X=0\} = q^n$
    $P\{X=1\} = C_n^1 pq^{n-1}$
    $P\{X=2\} = C_n^2 p^2 q^{n-2}$
    ......
    In general, for $k = 0,1,2,\ldots n$
    $P\{X=k\} = C_n^k p^k q^{n-k} = b(k;n,p)$
  - C.d.f:
    $$F(x) = P\{X \le x\} = \sum_{k=0}^{\lfloor x \rfloor} C_n^k p^k q^{n-k} = B(x;n,p)$$
  - E[X]: Binomial r.v. X can be represented as X=X1+X2+…+Xn, each Xi is a Bernoulli r.v. and they are independently identically distributed (i.i.d), and E[X]=E[X1]+E[X2]+…E[Xn]=*np*

Example Discrete *r.v.*s (3)

- Geometric r.v.: is a r.v. that counts the number of independent Bernoulli trials until the first success is encountered.

$$\Omega = \{0^{i-1}1 \mid i = 1,2,3\ldots\},\ \text{where}\quad 0: \text{failure};\quad 1: \text{success}$$

– *P.m.f*:

$$P\{X=0\}=p$$
$$P\{X=1\}=qp$$
$$P\{X=2\}=q^2 p$$
......

In general, for $k=0,1,2,\ldots$

$$P\{X=k\}=q^k p$$

– *C.d.f:* $F_X(x)=P\{X\le x\}=\sum_{k=0}^{\lfloor x\rfloor} q^k p$

– *E[X]:* q/p

Example Discrete *r.v.*s (4)

- Poisson r.v.: is a r.v. defined by its p.m.f:

$$P\{X=k\}=e^{-\lambda}\frac{\lambda^k}{k!},\ \lambda\,(\text{constant})>0;\ k=0,1,2,3\ldots\ldots$$

– *E[X]*= λ

– *Proposition:* If X, Y are independent Poisson *r.v.*s with parameter λ, β respectively. Then Z=X+Y is also a Poisson r.v. with parameter $\lambda+\beta$, i.e., the p.m.f. of Z is

$$P\{Z=k\}=e^{-(\lambda+\beta)}\frac{(\lambda+\beta)^k}{k!}$$

Poisson $r.v.$s (Cont'd)

- *Applications*: Poisson r.v.s are good for counting things like
  - # of transactions arriving to a DBMS per second
  - # of customers arriving at a bank in an hour
  - # of packets arriving at a switch per second
  - In studying queueing system, the # of job arriving, # of job completing service is usually able to be modeled as Possion.

Agenda

- Random Variable ($r.v.$)
  - Basic concepts
  - Discrete $r.v.s$
  - **Continuous $r.v.s$**

Related reading: Allen's Ch. 2.5~2.6, 3.0~3.2

Continuous Random Variable

✓ Definitions

– A r.v. that can take on a range of real values (uncountable!) is said to be a ***continuous*** r.v.

- <u>**Characteristic functions**</u>
  – **Probability density function (p.d.f.) for continuous *r.v.***
- Important parameters of *r.v.*s
  – Expectation/Mean: E[X], the *k*th moment
  – Variance/standard deviation
  – Squared coefficient of variation (C.O.V.) and C.O.V.
- Important example *r.v.*s
  – *Continuous*: Uniform, Normal, Exponential

Continuous *r.v.s* (1)

- <u>Probability density function (p.d.f.)</u>

The p.d.f. denoted by $f_X(x)$ for a continuous *r.v.* X is defined by

$$f_X(x) = F_X'(x) = \frac{dF_X(x)}{dx}$$

– $F_X(x)$ is *c.d.f* of r.v. X; since $F_X(x)$ is non-decreasing in x → $f_X(x) \geq 0$, $\forall\, x \in$ R

– $\forall\, x$, $P(X=x)= 0$; p.m.f. of a continuous r.v. assumes only the value ZERO!

– And

$$F_X(a) = P(X \leq a) = \int_{-\infty}^{a} f_X(x)dx$$

$$F_X(\infty) = \int_{-\infty}^{\infty} f_X(x)dx = 1$$

$$\int_a^b f_X(x)dx = P(a \leq X \leq b) = F_X(b) - F_X(a)$$

Continuous *r.v.s* (2)

- Expectation/Mean/Expected value of X

Let X be a continuous *r.v.* with p.d.f. $f_X(x)$, the mean of X, denoted by μ=E[X], is defined by

$$\mu = E[X] = \int_{-\infty}^{\infty} x f_X(x) dx$$

- Expectation of a function of a r.v. X

For any real-valued function *g:*

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$$

- The *k*th moment of X

The *k*th moment of X is defined by $E[X^k]$, $\{k = 1, 2, 3, \ldots\}$.

$$E[X^k] = \int_{-\infty}^{\infty} x^k f_X(x) dx$$

Continuous *r.v.s* (3)

- Variance of X

If X is a r.v. with mean μ, then the variance of X, denoted by Var(X)= δ^2, is defined by Var(X)= $E[(X-\mu)^2]$

– If X is a continuous *r.v.*, then

$$Var[X] = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x) dx$$

– An alternative formula:

$$Var[X] = E[X^2] - (E[X])^2 = E[X^2] - \mu^2$$

Proof:

$$\begin{aligned} Var[X] &= E[(X-\mu)^2] = E[X^2 - 2\mu X + \mu^2] \\ &= E[X^2] - 2\mu E[X] + E[\mu^2] \\ &= E[X^2] - 2\mu^2 + \mu^2 = E[X^2] - \mu^2 \end{aligned}$$

Continuous *r.v.s* (4)

- Standard deviation of X

  The squared root of the variance Var(X) is called the standard deviation of X:

$$\sigma = SD(X) = \sqrt{Var[X]}$$

- Squared coefficient of variation (C.O.V)

$$C_X^2 = \frac{Var[X]}{E[X]^2} = \frac{\sigma^2}{\mu^2}$$

- Coefficient of variation (C.O.V)

  is defined as the squared root of C_X^2:

$$COV = \sqrt{\frac{Var[X]}{E[X]^2}} = \frac{\sigma}{\mu}$$

Continuous Random Variable

- ✓ Definitions
  - ✓ A r.v. that can take on a range of real values (uncountable!) is said to be a *continuous* r.v.
- ✓ Characteristic functions
  - ✓ Probability density function (p.d.f.) for continuous *r.v.*
- ✓ Important parameters of *r.v.*s
  - ✓ Expectation/Mean: E[X], the *k*th moment
  - ✓ Variance/standard deviation
  - ✓ Squared coefficient of variation (C.O.V.) and C.O.V.
- **Important example *r.v.s***
  - *Continuous*: Uniform, Normal, Exponential

Example Continuous *r.v.s* (1)

- Uniform r.v.: a continuous r.v. is uniformly distributed on the interval a to b if it's p.d.f.

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \textbf{if } a < x < b \\ 0 & \textbf{otherwise} \end{cases}$$

  - *C.d.f: F(x)*

$$F(x) = \int_{-\infty}^{x} f(u)du = \begin{cases} 0 & x \le a \\ \dfrac{x-a}{b-a} & a < x < b \\ 1 & x \ge b \end{cases}$$

  - *E[X]*

$$E[X] = \frac{a+b}{2}$$

Example Continuous *r.v.s* (2)

- Normal / Gaussian r.v.: a continuous r.v. is normally distributed with parameter μ and σ^2 if its p.d.f:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

  - *Write as*: $X \sim N(\mu, \sigma^2)$
  - * *Pd.f.* of a normal r.v. has the well-known *bell-shaped* curve, that is symmetric about μ
  - Standard normal distribution: μ=0 and σ=1, $X \sim N(0,1)$
  - E[X] = μ, Var[X] = σ^2 (*Theorem 3.2.3*)
  - Suppose $X_1, X_2, \ldots, X_n$ are n independent Normal r.v.s such that $X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$, …, $X_n \sim N(\mu_n, \sigma_n^2)$, then $Y = X_1 + X_2 + \ldots + X_n$ is normally distributed with mean = $\mu_1 + \mu_2 + \ldots + \mu_n$ variance = $\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2$ (*Theorem 3.2.4*)

Example Continuous *r.v.s* (3)

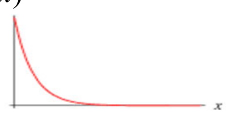
- Exponential r.v.: a continuous r.v. X has an exponential distribution with parameter λ if its *p.d.f*:

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

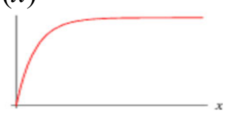
  - *C.d.f*:

$$F(x) = \int_{-\infty}^{x} f(x)dx = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

f(x)



F(x)



  - E[X]=1/λ
  - Var[X]=1/λ^2

Exponential *r.v.* (Cont'd)

- The *r*th percentile value of a *r.v.*: $\pi[r]$
  - $P\{X \leq \pi[r]\} = r/100$
  - For exponential r.v. with parameter λ: $\pi[r]$?

- Applications:
  - Time between 2 successive job arrivals to a file server (inter-arrival time)
  - Service time at a server in a queueing network
  - Time-to-failure of a component
  - Time required to repair a component that has malfunctioned

Properties of Exponential Distribution (1)

- "Memoryless/Markov" property:
  - The future is independent of the past!

$$P\{X > t+h \mid X > t\} = P\{X > h\} \quad \forall t, h > 0$$

Properties of Exponential Distribution (2)

- Suppose $X_1, X_2, \ldots, X_n$ are n independent exponential r.v.s with parameters $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectively, and $Y = min\{X_1, X_2, \ldots, X_n\}$. Then Y has an exponential distribution with parameter $\lambda = \lambda_1 + \lambda_2 + \ldots + \lambda_n$. (*Theorem 3.2.1 (h)*)
- If X is an exponential r.v. with parameter $\lambda>0$ →
  - $E[X] = 1/\lambda$
  - $Var[X] = 1/\lambda^2 = E[X]^2$
  - $E[X^k] = k!/\lambda^k = k!E[X]^k$
  - The rth percentile value $\pi[r]$ defined by $P\{X \le \pi[r]\} = r/100$ is given by $$\pi[r] = E[X]\ln\left(\frac{100}{100-r}\right)$$

Poisson and Exponential (1)

of arrivals of some entity (customer, job) per unit of time Y is *Poisson* distributed with parameter λ

$$\Updownarrow$$

Interarrival times T (time between any two successive arrivals) are independent, identically *exponentially* distributed with parameter λ

Poisson and Exponential (2)

- Specifically:
  - If # of arrivals, Y, of some entity per unit of time is described by a Poisson r.v. with parameter λ. Then the time T between any two successive arrivals (inter-arrival time) is independent of the inter-arrival time of any other successive arrivals and has an exponential distribution with parameter λ. Thus, $E[T]=1/\lambda$, and $P\{T \le t\}=1-\exp(-\lambda t)$ for $t \ge 0$ (*Theorem 3.2.1 (f)*)
  - Suppose inter-arrival times are i.i.d, exponential r.v.s, each with mean $1/\lambda$. Then the number of arrivals Y_t, in any interval of length $t > 0$, has a Poisson distribution with parameter λt, i.e., (*Theorem 3.2.1 (g)*)

$$P\{Y_t = k\} = e^{-\lambda t}\frac{(\lambda t)^k}{k!},\quad k = 0,1,2,3\ldots \quad \ldots$$

Hands-on Problem

- Personnel of the Farout Eng. Company use an online terminal to make routine engineering calculation. If the time each engineer spends in a session at a terminal has an exponential distribution with an average value of 36 minutes. Find
  - Probability that an engineer will spend 30 minutes or less at the terminal?
  - Probability that an engineer will use it for more than 1 hour?
  - Probability that an engineer will spend more than another 1 hour at the terminal if he or she has already been at the terminal for 1 hour?
  - 90% of the sessions end in less than R minutes. What is the value of R?

Next Topic

- Jointly distributed *r.v.s*

Things to Do

- Homework
- Project proposal due **February 23, Friday**