به نام خدا

Performance Evaluation of Computer Systems

Prof. Ali Movaghar

Fall 2022

Performance Modeling and Design of Computer Systems

3- PROBABILITY REVIEW

Probability Space

- (Ω, Σ, p)
- An event : $E \epsilon \Sigma$ is any subset of the sample space, Ω
- $|\Sigma| = 2^{|\Omega|}$
- $p : \Sigma \to [0\ 1]$

Probability Space

- $E1$ and $E2$ are mutually exclusive
  - $E1 \cap E2 \ = \emptyset$

- $E_1, E_2, \ldots, E_n$ are events such that
  - $\forall\ i, j\ \ E_i \cap E_j = \emptyset$
  - $\cup_{i=1}^{n} E_i \ = F$

  => Events $E_1, E_2, \ldots, E_n$ **partition** set F.

Conditional Probability & Independent Events

- Conditional probability of event E given event F :
  - $P\{E|F\}$
  - $P\{E|F\} = \frac{P\{E \cap F\}}{P\{F\}}$

- Events E and F are independent if :
  - $P\{E \cap F\} = P\{E\} \times P\{F\}$
  - $P\{E|F\} = \frac{P\{E \cap F\}}{P\{F\}} = \mathrm{P\{E\}}$

- Events E and F are conditionally independent given event G, where $P\{G\} > 0$, if :
  - $P\{E \cap F|G\} = P\{E|G\} \times P\{F|G\}$

Law of Total Probability

- Let $F_1, \ldots, F_n$ **partition** the state space Ω Then,
  - $P\{E\} = \sum_{i=1}^{n} P\{E \cap F_i\}$

$$= \sum_{i=1}^{n} P\{E|F_i\} \times P\{F_i\}$$

Bayes Law

- $P\{F|E\} = \frac{P\{E\cap F\}}{P\{E\}} = \frac{P\{E|F\}.P\{F\}}{P\{E\}}$
- **Extended Bayes Law**
  - Let $F, F_2, \dots, F_n$ **partition** the state space Ω. Then,
    - $P\{F|E\} = \frac{P\{E\cap F\}}{P\{E\}} = \frac{P\{E|F\}.P\{F\}}{P\{E\}} = \frac{P\{E|F\}.P\{F\}}{\sum_{i=1}^{n} P\{E|F_i\}\times P\{F_i\}}$

Random Variable

- $X : \Omega \rightarrow R$
  - Real-valued function of the outcome of an experiment
  - All the theorems that we learned about events apply to random variables as well
  - e.g. Total probability

Probabilities and Densities : Discrete

- Probability mass function (**p.m.f.)**
  - $P_X(a) = P\{X = a\},\ where \sum_x P_X(a) = 1$
- Cumulative distribution function
  - $F_X(a) = P\{X \leqslant a\} = \sum_{x \leqslant a} P_X(x)$
  - $\overline{F_X}(a) = P\{X > a\} = \sum_{x > a} P_X(x) = 1 - F_X(a)$
- Samples
  - Bernoulli (p)
  - Binomial (n, p)
  - Geometric (p)
  - Poisson (λ)

Probabilities and Densities : Continuous

- Probability density function (**p.d.f.**)
  - $P(a \leqslant X \leqslant b) = \int_a^b f_X(x)dx\,,\ and\ where\ \int_{-\infty}^{\infty} f_X(x)dx = 1$
  - $f_X(x) \neq P\{X = x\}$
  - $f_X(x)dx = P\{x \le X \le x + dx\}$
- Cumulative distribution function
  - $F_X(a) = P\{-\infty \le X \le a\} = \int_{-\infty}^{a} f_X(x)dx$
  - $\overline{F_X}(a) = P\{X > a\} = 1 - F_X(a)$
  - $f_X(x) = \frac{d}{dx}\int_{-\infty}^{x} f(t)dt = \frac{d}{dx}F_X(x)$
- Samples
  - Uniform (a, b)
  - Exp (λ)
  - Pareto (α)

Expectation and Variance

- Discrete random variable : X
  - $E[X] = \sum_x x . p_X(x)$
  - $E[X^i] = \sum_x x^i . p_X(x)$

- Continuous random variable : X
  - $E[X] = \int_{-\infty}^{\infty} x . f_X(x) dx$
  - $E[X^i] = \int_{-\infty}^{\infty} x^i . f_X(x) dx$

Expectation of a Function

- Discrete random variable : X
  - $E[g(X)] = \sum_x g(x).\, p_X(x)$

- Continuous random variable : X
  - $E[g(X)] = \int_{-\infty}^{\infty} g(x).\, f_X(x)dx$

Variance

- Expected squared difference of X from its mean
  - $Var(X) = E[(X - E[X])^2]$
  - $Var(X) = E[X^2] - (E(X))^2$

Table 3.2. *Discrete and continuous distributions*

Distribution	p.m.f. $p_X(x)$	Mean	Variance
Bernoulli(p)	$p_X(0) = 1 - p\,;\, p_X(1) = p$	p	$p(1-p)$
Binomial(n, p)	$p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$, $x = 0, 1, \ldots, n$	np	$np(1-p)$
Geometric(p)	$p_X(x) = (1-p)^{x-1} p, \quad x = 1, 2, \ldots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Poisson(λ)	$p_X(x) = e^{-\lambda} \cdot \frac{\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots$	λ	λ
Distribution	**p.d.f. $f_X(x)$**	**Mean**	**Variance**
Exp(λ)	$f_X(x) = \lambda e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Uniform(a, b)	$f_X(x) = \frac{1}{b-a}, \quad \text{if } a \le x \le b$	$\frac{b+a}{2}$	$\frac{(b-a)^2}{12}$
Pareto(α), $0 < \alpha < 2$	$f_X(x) = \alpha x^{-\alpha-1}, \quad \text{if } x > 1$	$\begin{cases} \infty & \text{if } \alpha \le 1 \\ \frac{\alpha}{\alpha-1} & \text{if } \alpha > 1 \end{cases}$	∞
Normal(μ, σ^2)	$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$, $-\infty < x < \infty$	μ	σ^2

Joint Probabilities and Independence

- Discrete random variables X and Y
- Joint probability mass function
  - $p_{X,Y}(x,y) = P\{X = x \,\&\, Y = y\}$
  - $p_X(x) = \sum_y p_{X,Y}(x,y)$
  - $p_Y(y) = \sum_x p_{X,Y}(x,y)$
- X and Y are independent
  - $X \perp Y$
  - $p_{X,Y}(x,y) = p_X(x).\, p_Y(y)$

Joint Probabilities and Independence

- Contenous random variables X and Y
- Joint probability density function
  - $\int_{c}^{d}\int_{a}^{b} f_{X,Y}\,(x, y) = \boldsymbol{P}\{a < X < b \,\&\, c < Y < d\}$
  - $f_X\,(x) = \int_{-\infty}^{\infty} f_{X,Y}\,(x, y)dy$
  - $f_Y\,(y) = \int_{-\infty}^{\infty} f_{X,Y}\,(x, y)dx$
- X and Y are independent
  - $X \perp Y$
  - $f_{X,Y}\,(x, y) = f_x(x).\, f_y(y), \forall x, y$

Theorem 3.20

Theorem 3.20 *If* $X \perp Y$, *then* $\mathbf{E}[XY] = \mathbf{E}[X] \cdot \mathbf{E}[Y]$.

Proof

$$\begin{aligned}
\mathbf{E}[XY] &= \sum_x \sum_y xy \cdot \mathbf{P}\{X = x, Y = y\} \\
&= \sum_x \sum_y xy \cdot \mathbf{P}\{X = x\}\mathbf{P}\{Y = y\} \quad \text{(by definition of } \perp\text{)} \\
&= \sum_x x\mathbf{P}\{X = x\} \cdot \sum_y y\mathbf{P}\{Y = y\} \\
&= \mathbf{E}[X]\,\mathbf{E}[Y]
\end{aligned}$$

The same argument works for continuous r.v.'s. ■

- We also have,
  - $E[g(X)f(Y)] = E[g(X)] \times E[f(Y)]$

Question?

- $E[XY] = E[X]E[Y] \overset{?}{\Rightarrow} X \perp Y$
- No, see Exercise3.10

Conditional Probabilities : Discrete

- Conditional $\boldsymbol{p.m.f}$. of X given event A
  - $p_{X|A}(x) = P\{X = x \mid A\} = \frac{P\{(X=x)\cap A\}}{P\{A\}}$
- Conditional expectation of X given event A
  - $\mathrm{E}(X|A) = \sum_x x.\,p_{X|A}(x) = \sum_x x.\,\frac{P\{(X=x)\cap A\}}{P\{A\}}$

Conditional Probabilities : Continuous

- Conditional $\boldsymbol{p}.\boldsymbol{d}.\boldsymbol{f}.$ of X given event A
  - $f_{X|A}(x) = \begin{cases} \frac{f_X(x)}{P\{X \epsilon A\}}, & if\ x \in A \\ 0, & otherwise \end{cases}$

- Conditional expectation of X given event A
  - $E(X|A) = \int_{-\infty}^{\infty} x f_{X|A}(x) dx$

    $= \int_A x f_{X|A}(x) dx = \frac{1}{\boldsymbol{P}\{X \epsilon A\}} \int_A x f_X(x) dx$

Probabilities and Expectations via Conditioning

- Let $F, F_2, \ldots, F_n$ **partition** the state space Ω
  - $\boldsymbol{P}\{E\} = \sum_{i=1}^{n} \boldsymbol{P}\{E|F_i\}\boldsymbol{P}\{F_i\}$
- Law of Total Probability for Discrete Random Variables
  - $\boldsymbol{P}\{X = k\} = \sum_y \boldsymbol{P}\{X = k \mid Y = y\} . \boldsymbol{P}\{Y = y\}$

Theorem 3.25 *For discrete random variables,*

$$\mathbf{E}[X] = \sum_y \mathbf{E}[X \mid Y = y]\,\mathbf{P}\{Y = y\}.$$

Similarly for continuous random variables,

$$\mathbf{E}[X] = \int \mathbf{E}[X \mid Y = y]\, f_Y(y)dy.$$

- $E[g(X)] = \sum_y E[g(X)|Y = y] . P\{Y = y\}$

Linearity of Expectation

Theorem 3.26 (Linearity of Expectation) *For random variables X and Y,*

$$\mathbf{E}[X+Y] = \mathbf{E}[X] + \mathbf{E}[Y].$$

Proof Here is a proof in the case where X and Y are continuous. The discrete case is similar: Just replace $f_{X,Y}(x,y)$ with $p_{X,Y}(x,y)$.

$$\begin{aligned}
&\mathbf{E}[X+Y] \\
&= \int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty}(x+y)f_{X,Y}(x,y)dxdy \\
&= \int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty}xf_{X,Y}(x,y)dxdy + \int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty}yf_{X,Y}(x,y)dxdy \\
&= \int_{x=-\infty}^{\infty}\int_{y=-\infty}^{\infty}xf_{X,Y}(x,y)dydx + \int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty}yf_{X,Y}(x,y)dxdy \\
&= \int_{x=-\infty}^{\infty}x\int_{y=-\infty}^{\infty}f_{X,Y}(x,y)dydx + \int_{y=-\infty}^{\infty}y\int_{x=-\infty}^{\infty}f_{X,Y}(x,y)dxdy \\
&= \int_{x=-\infty}^{\infty}xf_X(x)dx + \int_{y=-\infty}^{\infty}yf_Y(y)dy \\
&= \mathbf{E}[X] + \mathbf{E}[Y] \qquad \blacksquare
\end{aligned}$$

Linearity of Expectation

Theorem 3.27 *Let X and Y be random variables where $X \perp Y$. Then*

$$\mathbf{Var}(X+Y) = \mathbf{Var}(X) + \mathbf{Var}(Y).$$

Proof

$$\begin{aligned}
\mathbf{Var}(X+Y) &= \mathbf{E}\left[(X+Y)^2\right] - (\mathbf{E}[(X+Y)])^2 \\
&= \mathbf{E}\left[X^2\right] + \mathbf{E}\left[Y^2\right] + 2\mathbf{E}[XY] \\
&\quad -(\mathbf{E}[X])^2 - (\mathbf{E}[Y])^2 - 2\mathbf{E}[X]\,\mathbf{E}[Y] \\
&= \mathbf{Var}(X) + \mathbf{Var}(Y) \\
&\quad \underbrace{+2\mathbf{E}[XY] - 2\mathbf{E}[X]\,\mathbf{E}[Y]}_{\text{equals 0 if X} \perp \text{Y}}
\end{aligned}$$

Normal Distribution

Definition 3.28 A continuous r.v. X is said to be ***Normal***(μ, σ^2) or ***Gaussian***(μ, σ^2) if it has p.d.f. $f_X(x)$ of the form

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \qquad -\infty < x < \infty$$

where $\sigma > 0$. The parameter μ is called the ***mean***, and the parameter σ is called the ***standard deviation***.

Definition 3.29 A Normal$(0, 1)$ r.v. Y is said to be a ***standard Normal***. Its c.d.f. is denoted by

$$\Phi(y) = F_Y(y) = \mathbf{P}\{Y \leq y\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-t^2/2} dt.$$

Theorem 3.30 *Let* $X \sim$ *Normal*(μ, σ^2)*, then* $\mathbf{E}[X] = \mu$ *and* $\mathbf{Var}(X) = \sigma^2$.

Linear Transformation Property

Theorem 3.31 (Linear Transformation Property) *Let $X \sim Normal(\mu, \sigma^2)$. Let $Y = aX + b$, where $a > 0$ and b are scalars. Then $Y \sim Normal(a\mu + b, a^2\sigma^2)$.*

- Thus,
- $X \sim Normal(\mu, \sigma^2) \Leftrightarrow Y = \frac{X-\mu}{\sigma} \sim Normal(0, 1)$
- $P\{X < K\} = P\left\{\frac{X-\mu}{\sigma} < \frac{k-\mu}{\sigma}\right\} = P\left\{Y < \frac{k-\mu}{\sigma}\right\} = \Phi(\frac{k-\mu}{\sigma})$

Linear Transformation Property

Theorem 3.32 *If $X \sim Normal(\mu, \sigma^2)$, then the probability that X deviates from its mean by less than k standard deviations is the same as the probability that the standard Normal deviates from its mean by less than k.*

Proof Let $Y \sim \text{Normal}(0, 1)$. Then,

$$\mathbf{P}\{-k\sigma < X - \mu < k\sigma\} = \mathbf{P}\left\{-k < \frac{X - \mu}{\sigma} < k\right\} = \mathbf{P}\{-k < Y < k\} \qquad \blacksquare$$

Theorem 3.32 illustrates why it is often easier to think in terms of standard deviations than in absolute values.

Central Limit Theorem

Theorem 3.33 (Central Limit Theorem (CLT)) *Let $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d. r.v.'s with common mean μ and variance σ^2, and define*

$$Z_n = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}.$$

Then the c.d.f. of Z_n converges to the standard normal c.d.f.; that is,

$$\lim_{n\to\infty} \mathbf{P}\{Z_n \le z\} = \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx$$

for every z.

- $\boldsymbol{Binomial(n, p)}$distribution, which is a sum of $\boldsymbol{i.i.d.}$ $\boldsymbol{Bernoulli(p)}$ r.v.'s, converges to a Normal distribution when $\boldsymbol{n}$ is high.
- $\boldsymbol{Poisson(\lambda)}$ distribution is also well approximated by a $\boldsymbol{Normal}$ distribution with mean $\boldsymbol{\lambda}$ and variance $\boldsymbol{\lambda}$.

Sum of a Random Number of Random Variables

- Number of these variables is itself a random variable
- $\boldsymbol{N}$: non-negative integer-valued random variable

- $S = \sum_{i=1}^{N} X_i, N \perp X_i$

- $E[S], E[S^2], \ldots$?
  - Linearity equations only apply when $\boldsymbol{N}$ is a constant
  - ?

$E[S]$

- Condition on the value of N, and then apply linearity of expectation

$$
\begin{aligned}
\mathbf{E}[S] = \mathbf{E}\left[\sum_{i=1}^{N} X_i\right] &= \sum_n \mathbf{E}\left[\sum_{i=1}^{N} X_i \,\middle|\, N = n\right] \cdot \mathbf{P}\{N = n\} \\
&= \sum_n \mathbf{E}\left[\sum_{i=1}^{n} X_i\right] \cdot \mathbf{P}\{N = n\} \\
&= \sum_n n\mathbf{E}[X] \cdot \mathbf{P}\{N = n\} \\
&= \mathbf{E}[X] \cdot \mathbf{E}[N]
\end{aligned}
$$

$E[S^2]$

- ?

Theorem 3.34 *Let $X_1, X_2, X_3, \ldots$ be i.i.d. random variables. Let*

$$S = \sum_{i=1}^{N} X_i, \quad N \perp X_i.$$

Then

$$\begin{aligned}
\mathbf{E}[S] &= \mathbf{E}[N]\,\mathbf{E}[X], \\
\mathbf{E}\left[S^2\right] &= \mathbf{E}[N]\,\mathbf{Var}(X) + \mathbf{E}\left[N^2\right](\mathbf{E}[X])^2, \\
\mathbf{Var}(S) &= \mathbf{E}[N]\,\mathbf{Var}(X) + \mathbf{Var}(N)\,(\mathbf{E}[X])^2.
\end{aligned}$$