Performance Evaluation of Computer Systems

Dr. Ali Movaghar

Fall 2014
3- PROBABILITY REVIEW
Probability Space

• \((\Omega, \Sigma, p)\)
• An event \(E \in \Sigma\) is any subset of the sample space, \(\Omega\)
• \(|\Sigma| = 2^{|\Omega|}\)
• \(p: \Sigma \rightarrow [0, 1]\)
Probability Space

• $E_1$ and $E_2$ are mutually exclusive
  – $E_1 \cap E_2 = \emptyset$

• $E_1, E_2, \ldots, E_n$ are events such that
  – $\forall i, j \ E_i \cap E_j = \emptyset$
  – $\bigcup_{i=1}^{n} E_i = F$
  => 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:
  - $P\{E \cap F\} = P\{E\} \times P\{F\}$
  - $P\{E|F\} = \frac{P\{E \cap F\}}{P\{F\}} = P\{E\}$

- 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 $\Omega$. 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

- \( P\{F|E\} = \frac{P\{E \cap F\}}{P\{E\}} = \frac{P\{E|F\} \cdot P\{F\}}{P\{E\}} \)

- **Extended Bayes Law**
  
  - Let \( F, F_2, \ldots, F_n \) partition the state space \( \Omega \). Then,

    \[
    P\{F|E\} = \frac{P\{E \cap F\}}{P\{E\}} = \frac{P\{E|F\} \cdot P\{F\}}{P\{E\}} = \frac{P\{E|F\} \cdot P\{F\}}{\sum_{i=1}^{n} P\{E|F_i\} \times P\{F_i\}}
    \]
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**
  
  - $F_X(a) = P\{X \leq a\} = \sum_{x \leq a} P_X(x)$
  
  - $\overline{F_X}(a) = P\{X > a\} = \sum_{x > a} P_X(x) = 1 - F_X(a)$

- **Samples**
  
  - Bernoulli ($p$)
  
  - Binomial ($n, p$)
  
  - Geometric ($p$)
  
  - Poisson ($\lambda$)
Probabilities and Densities: Continuous

- **Probability density function (p.d.f.)**
  - \( P(a \leq X \leq b) = \int_a^b f_X(x)dx \), and where \( \int_{-\infty}^{\infty} f_X(x)dx = 1 \)
  - \( f_X(x) \neq P\{X = x\} \)
  - \( f_X(x)dx = P\{x \leq X \leq x + dx\} \)

- **Cumulative distribution function**
  - \( F_X(a) = P\{-\infty \leq X \leq a\} = \int_{-\infty}^{a} f_X(x)dx \)
  - \( F_X(a) = P\{X > a\} = 1 - F_X(a) \)
  - \( f_X(x) = \frac{d}{dx} \int_{-\infty}^{x} f(t)dt = \frac{d}{dx} F_X(x) \)

- **Samples**
  - Uniform \((a, b)\)
  - Exp \((\lambda)\)
  - Pareto \((\alpha)\)
Expectation and Variance

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

• Continuous random variable : X
  - $E[X] = \int_{-\infty}^{\infty} x \cdot f_X(x) dx$
  - $E[X^i] = \int_{-\infty}^{\infty} x^i \cdot f_X(x) dx$
Expectation of a Function

• Discrete random variable: \( X \)
  \[ E[g(X)] = \sum_x g(x) \cdot p_X(x) \]

• Continuous random variable: \( X \)
  \[ E[g(X)] = \int_{-\infty}^{\infty} g(x) \cdot f_X(x) \, dx \]
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>
<thead>
<tr>
<th>Distribution</th>
<th>p.m.f. $p_X(x)$</th>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bernoulli($p$)</td>
<td>$p_X(0) = 1 - p$; $p_X(1) = p$</td>
<td>$p$</td>
<td>$p(1 - p)$</td>
</tr>
<tr>
<td>Binomial($n, p$)</td>
<td>$p_X(x) = \binom{n}{x}p^x (1-p)^{n-x}$, $x = 0, 1, \ldots, n$</td>
<td>$np$</td>
<td>$np(1 - p)$</td>
</tr>
<tr>
<td>Geometric($p$)</td>
<td>$p_X(x) = (1 - p)^{x-1}p$, $x = 1, 2, \ldots$</td>
<td>$\frac{1}{p}$</td>
<td>$\frac{1-p}{p^2}$</td>
</tr>
<tr>
<td>Poisson($\lambda$)</td>
<td>$p_X(x) = e^{-\lambda} \frac{\lambda^x}{x!}$, $x = 0, 1, 2, \ldots$</td>
<td>$\lambda$</td>
<td>$\lambda$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Distribution</th>
<th>p.d.f. $f_X(x)$</th>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp($\lambda$)</td>
<td>$f_X(x) = \lambda e^{-\lambda x}$</td>
<td>$\frac{1}{\lambda}$</td>
<td>$\frac{1}{\lambda^2}$</td>
</tr>
<tr>
<td>Uniform($a, b$)</td>
<td>$f_X(x) = \frac{1}{b-a}$, if $a \leq x \leq b$</td>
<td>$\frac{b+a}{2}$</td>
<td>$\frac{(b-a)^2}{12}$</td>
</tr>
<tr>
<td>Pareto($\alpha$), $0 &lt; \alpha &lt; 2$</td>
<td>$f_X(x) = \alpha x^{-\alpha - 1}$, if $x &gt; 1$</td>
<td>$\begin{cases} \infty &amp; \text{if } \alpha \leq 1 \ \frac{\alpha}{\alpha - 1} &amp; \text{if } \alpha &gt; 1 \end{cases}$</td>
<td>$\infty$</td>
</tr>
<tr>
<td>Normal($\mu, \sigma^2$)</td>
<td>$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2}$, $-\infty &lt; x &lt; \infty$</td>
<td>$\mu$</td>
<td>$\sigma^2$</td>
</tr>
</tbody>
</table>
Joint Probabilities and Independence

• Discrete random variables $X$ and $Y$

• Joint probability mass function
  
  $\begin{align*}
  & - p_{X,Y}(x, y) = P\{X = x \& Y = y\} \\
  & - p_X(x) = \sum_y p_{X,Y}(x, y) \\
  & - p_Y(y) = \sum_x p_{X,Y}(x, y)
  \end{align*}$

• $X$ and $Y$ are independent
  
  $\begin{align*}
  & - X \perp Y \\
  & - p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y)
  \end{align*}$
Joint Probabilities and Independence

- Contenous random variables $X$ and $Y$
- Joint probability density function
  \[ - \int_c^d \int_a^b f_{X,Y} (x, y) = P\{a < X < b \& c < Y < d\} \]
  \[ - f_X (x) = \int_{-\infty}^{\infty} f_{X,Y} (x, y) dy \]
  \[ - f_Y (y) = \int_{-\infty}^{\infty} f_{X,Y} (x, y) dx \]
- $X$ and $Y$ are independent
  \[ - X \perp Y \]
  \[ - f_{X,Y} (x, y) = f_x(x).f_y(y), \forall x, y \]
Theorem 3.20

If $X \perp Y$, then $E[XY] = E[X] \cdot E[Y]$.

Proof

$$E[XY] = \sum_x \sum_y xy \cdot P\{X = x, Y = y\}$$

$$= \sum_x \sum_y xy \cdot P\{X = x\} P\{Y = y\} \quad \text{(by definition of } \perp)$$

$$= \sum_x x P\{X = x\} \cdot \sum_y y P\{Y = y\}$$

$$= E[X] E[Y]$$

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?

• $E[XY] = E[X]E[Y]$? $\Rightarrow X \perp Y$

• No, see Exercise 3.10
Conditional Probabilities: Discrete

- Conditional probability mass function of $X$ given event $A$

$$p_{X \mid A}(x) = P\{X = x \mid A\} = \frac{P\{(X=x) \cap A\}}{P\{A\}}$$

- Conditional expectation of $X$ given event $A$

$$E(X \mid A) = \sum_x x \cdot p_{X \mid A}(x) = \sum_x x \cdot \frac{P\{(X=x) \cap A\}}{P\{A\}}$$
Conditional Probabilities : Continuous

• Conditional *p. d. f.* of $X$ given event $A$

  $$f_{X|A}(x) = \begin{cases} \frac{f_X(x)}{P\{X \in A\}}, & \text{if } x \in A \\ 0, & \text{otherwise} \end{cases}$$

• Conditional expectation of $X$ given event $A$

  $$E(X|A) = \int_{-\infty}^{\infty} xf_{X|A}(x)\,dx$$

  $$= \int_{A} xf_{X|A}(x)\,dx = \frac{1}{P\{X \in A\}} \int_{A} xf_{X}(x)\,dx$$
Probabilities and Expectations via Conditioning

- Let $F, F_2, \ldots, F_n$ partition the state space $\Omega$
  \[
  P\{E\} = \sum_{i=1}^{n} P\{E | F_i\} P\{F_i\}
  \]

- Law of Total Probability for Discrete Random Variables
  \[
  P\{X = k\} = \sum_y P\{X = k | Y = y\} . P\{Y = y\}
  \]

**Theorem 3.25** For discrete random variables,

\[
E[X] = \sum_y E[X | Y = y] P\{Y = y\}.
\]

Similarly for continuous random variables,

\[
E[X] = \int E[X | Y = y] f_Y(y) dy.
\]

- $E[g(X)] = \sum_y E[g(X)|Y = y]. P\{Y = y\}$
Theorem 3.26 (Linearity of Expectation) For random variables $X$ and $Y$,

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

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 $p_{X,Y}(x, y)$.

$$\begin{align*}
\mathbb{E}[X + Y] &= \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{\infty} (x + y) f_{X,Y}(x, y) \, dx \, dy \\
&= \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{\infty} x f_{X,Y}(x, y) \, dx \, dy + \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{\infty} y f_{X,Y}(x, y) \, dx \, dy \\
&= \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} x f_{X,Y}(x, y) \, dy \, dx + \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{\infty} y f_{X,Y}(x, y) \, dx \, dy \\
&= \int_{x=-\infty}^{\infty} x \int_{y=-\infty}^{\infty} f_{X,Y}(x, y) \, dy \, dx + \int_{y=-\infty}^{\infty} y \int_{x=-\infty}^{\infty} f_{X,Y}(x, y) \, dx \, dy \\
&= \int_{x=-\infty}^{\infty} x f_X(x) \, dx + \int_{y=-\infty}^{\infty} y f_Y(y) \, dy \\
&= \mathbb{E}[X] + \mathbb{E}[Y]
\end{align*}$$
Linearity of Expectation

**Theorem 3.27** Let $X$ and $Y$ be random variables where $X \perp Y$. Then
\[
\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).
\]

**Proof**
\[
\begin{align*}
\text{Var}(X + Y) &= \mathbb{E}[(X + Y)^2] - (\mathbb{E}[(X + Y)])^2 \\
&= \mathbb{E}[X^2] + \mathbb{E}[Y^2] + 2\mathbb{E}[XY] \\
&\quad - (\mathbb{E}[X])^2 - (\mathbb{E}[Y])^2 - 2\mathbb{E}[X] \mathbb{E}[Y] \\
&= \text{Var}(X) + \text{Var}(Y) \\
&\quad + 2\mathbb{E}[XY] - 2\mathbb{E}[X] \mathbb{E}[Y] \\
&\quad \text{equals 0 if } X \perp Y
\end{align*}
\]
Normal Distribution

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

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

where $\sigma > 0$. The parameter $\mu$ is called the *mean*, and the parameter $\sigma$ 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

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

**Theorem 3.30** Let $X \sim \text{Normal}(\mu, \sigma^2)$, then $E[X] = \mu$ and $\text{Var}(X) = \sigma^2$. 
Linear Transformation Property

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

- Thus,
- $X \sim \text{Normal}(\mu, \sigma^2) \iff Y = \frac{X-\mu}{\sigma} \sim \text{Normal}(0, 1)$
- $P\{X < K\} = P\left\{\frac{X-\mu}{\sigma} < \frac{k-\mu}{\sigma}\right\} = P\{Y < \frac{k-\mu}{\sigma}\} = \Phi\left(\frac{k-\mu}{\sigma}\right)$
Theorem 3.32  If $X \sim \text{Normal}(\mu, \sigma^2)$, 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,

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

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 $\mu$ and variance $\sigma^2$, 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,

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

for every $z$.

- **Binomial**($n, p$) distribution, which is a sum of i.i.d. **Bernoulli**($p$) r.v.’s, converges to a Normal distribution when $n$ is high.
- **Poisson**($\lambda$) distribution is also well approximated by a **Normal** distribution with mean $\lambda$ and variance $\lambda$. 
Sum of a Random Number of Random Variables

• Number of these variables is itself a random variable
• $N$: 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 $N$ is a constant
  – ?
\[ E[S] \]

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

\[
E[S] = E\left[ \sum_{i=1}^{N} X_i \right] = \sum_{n} E\left[ \sum_{i=1}^{N} X_i \mid N = n \right] \cdot P\{N = n\}
\]

\[
= \sum_{n} E\left[ \sum_{i=1}^{n} X_i \right] \cdot P\{N = n\}
\]

\[
= \sum_{n} nE[X] \cdot P\{N = n\}
\]

\[ = E[X] \cdot E[N] \]
$E[S^2]$
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

$$\mathbb{E}[s] = \mathbb{E}[N] \mathbb{E}[X],$$
$$\mathbb{E}[s^2] = \mathbb{E}[N] \text{Var}(X) + \mathbb{E}[N^2] (\mathbb{E}[X])^2,$$
$$\text{Var}(s) = \mathbb{E}[N] \text{Var}(X) + \text{Var}(N) (\mathbb{E}[X])^2.$$