خلاصه پایانترم شبیه‌سازی

فصل پنج

جمعیت هدف: جمعیتی که افراد از بین آن وارد سیستم می‌شوند.

ظرفیت سیستم: محدودیت سیستم برای تعداد افراد داخل آن

در سیستم‌هایی که جمعیت هدف نامحدود است فرض می‌کنیم سیستم هیچ‌گاه بیکار نیست.

شرایط pending: شرایطی که کاربر کاملا بیرون از سیستم و صف قرار دارد.

زمان runtime: مدت زمان بین دو مراجعه‌ی کاربر به سیستم. طول زمان شرایط pending.

رفتار صف: رفتاری که کاربران در صف از خود نشان می‌دهند. می‌تواند Balk (عدم ورود به صف طولانی)، Renege (خروج از صف کند) و Jockey (ورود به یک صف کوتاه‌تر) باشد.

سرویس‌دهی دسته‌ای (batch service): اگر یک سرور بتواند به طور همزمان به چند مشتری خدمت‌رسانی کند، سیستم دارای سرویس‌دهی دسته‌ای خواهد بود.

نام گذاری A/B/c/N/K: توزیع زمان بین ورود، توزیع زمان سرویس‌دهی، تعداد سرورهای داخل سیستم، ظرفیت سیستم و اندازه‌ی جمعیت هدف. (M: نمایی، D: قطعی، E_k: ارلانگ درجه k، PH: توزیع فاز، H: فرانمایی، G: هر توزیع عمومی)

$$L = \lim_{T \to \infty} \frac{1}{T} \int_0^T L(t) dt$$

$$Little'law\ (convservation\ equation): L = \lambda w$$

$$w = \frac{1}{\mu}\ \ (stable\ systems, G/G/1/\infty/\infty) \Longrightarrow w = \frac{\lambda}{\mu}$$

$$\rho = \frac{\lambda}{c\mu}$$

$$customer\ cost = \lambda w_Q \times unit\ cost$$

$$server\ cost = unit\ cost \times c(2-\rho)$$

برای صف‌های M/G/1:

$$P_0 = 1-\rho$$

$$L = \rho + \frac{\rho^2(1+\sigma^2\mu^2)}{2(1-\rho)}, \qquad L_Q = \frac{\rho^2(1+\sigma^2\mu^2)}{2(1-\rho)} = \left(\frac{\rho^2}{1-\rho}\right)\left(\frac{1+(cv)^2}{2}\right)$$

$$w = \frac{1}{\mu} + \frac{\lambda\left(\frac{1}{\mu^2}+\sigma^2\right)}{2(1-\rho)}, \qquad w_Q = \frac{\lambda\left(\frac{1}{\mu^2}+\sigma^2\right)}{2(1-\rho)}$$

برای صف‌های M/M/1 کافیست در معادلات فوق به جای σ^2 مقدار $\frac{1}{\mu^2}$ را جایگزین کنیم.

$$P_n = (1-\rho)\rho^n$$

$$L = \frac{\lambda}{\mu-\lambda} = \frac{\rho}{1-\rho}, \qquad L_Q = \frac{\lambda^2}{\mu(\mu-\lambda)} = \frac{\rho^2}{1-\rho}$$

$$w = \frac{1}{\mu-\lambda} = \frac{1}{\mu(1-\rho)}, \qquad w_Q = \frac{\lambda}{\mu(\mu-\lambda)} = \frac{\rho}{\mu(1-\rho)}$$

برای کاهش طول صف‌ها می‌توان ρ یا σ^2 (میزان تغییرات زمان سرویس) را کاهش داد.

برای صف‌هایی با چند سرور:

$$P_0 = \left\{\left[\sum_{n=0}^{c-1}\frac{\left(\frac{\lambda}{\mu}\right)^n}{n!}\right] + \left[\left(\frac{\lambda}{\mu}\right)^c\left(\frac{1}{c!}\right)\left(\frac{c\mu}{c\mu-\lambda}\right)\right]\right\}^{-1}$$

$$L = c\rho + \frac{(c\rho)^{c+1}P_0}{c(c!)(1-\rho)^2} = c\rho + \frac{\rho P(L(\infty)\geq c)}{1-\rho}$$

برای سیستم‌هایی با جامعه‌ی هدف محدود:

$$P_0 = \left\{ \sum_{n=0}^{c-1} \binom{K}{n} \left(\frac{\lambda}{\mu}\right)^n + \sum_{n=c}^{K} \frac{K!}{(K-n)!\,c!\,c^{n-c}} \left(\frac{\lambda}{\mu}\right)^n \right\}^{-1}$$

$$P_n = \begin{cases} \binom{K}{n} \left(\frac{\lambda}{\mu}\right)^n P_0, & n = 0,1,\dots,c-1 \\ \frac{K!}{(K-n)!\,c!\,c^{n-c}} \left(\frac{\lambda}{\mu}\right)^n, & n = c, c+1, \dots, K \end{cases}$$

$$L = \sum_{n=0}^{K} nP_n\,, \qquad w = \frac{L}{\lambda_e}, \qquad \rho = \frac{\lambda_e}{c\mu}$$

$$\lambda_e = \sum_{n=0}^{K} (K-n)\lambda P_n$$

در شبکه‌ی صف‌ها:

$$arrival\ rate\ i\ to\ j: \lambda_i P_{i,j}, \qquad P: fraction\ of\ customers\ that\ go\ from\ i\ to\ j$$

$$\lambda_i = a_j + \sum_{all\ i} \lambda_i p_{ij}$$

فصل شش

ویژگی‌های مهم اعداد تصادفی:

- یکنواختی: احتمال تولید هر کدام از آن‌ها با یکدیگر برابر باشد.
- استقلال: تولید یک عدد روی اعداد دیگر اثر نگذارد.

روش Linear Congruential (LCM)

از فرمول زیر استفاده می‌کنیم

$$X_i + 1 \ = \ (aX_i \ + \ c) \ mod\ m$$

$$R_i = \frac{X_i}{m}$$

که اعداد خروجی همان R_i ها هستند که اعدادی بین صفر تا یک هستند. مقادیر x و X_0 و c و m شانسی انتخاب می‌شوند.

برای رسیدن به طولانی‌ترین طول دوره دو حالت داریم:

$$c \neq 0, m = 2^b, \gcd(c, m) = 1, a = 1 + 4k \Rightarrow P = m$$
$$c = 0, X_0 \ \% \ 2 = 1, m = 2^b, a = 3 + 8k \mid 5 + 8k, \Rightarrow P = \frac{m}{4}$$
$$m \ prime, a^k - 1 \ \% \ m = 0 \Rightarrow P = m - 1$$

روش ترکیبی به صورت زیر است:

$$X_i = \left(\sum_{j=1}^{k} (-1)^{j-1} X_{i,j} \right) mod \ (m_1 - 1)$$

$$P = \frac{(m_1 - 1)(m_2 - 1) \dots (m_k - 1)}{2^{k-1}}$$

تست اعداد تصادفی:

باید ویژگی‌های یکنواختی و استقلال تست شوند. در هر دو حالت اگر نتوانیم فرض پوچ (null hypothesis) را رد کنیم، یعنی شواهد کافی نداریم و باید تست‌های بیشتری انجام شود.

Level of significance: $\alpha = P(Reject\ H0 \mid H0\ is\ True)$

تست کولموگروف: به صورت زیر تست را انجام می‌دهیم:

1. اعداد تولید شده را به صورت صعودی مرتب می‌کنیم.
2. مقدار $\frac{i}{N}$ را محاسبه می‌کنیم که همان مقدار $S_N(x)$ می‌باشد.
3. دو تفاضل $\frac{i}{N} - R_i$ و $R_i - \frac{i-1}{N}$ را محاسبه می‌کنیم و ماکسیمم مقادیر آن‌ها را بدست می‌آوریم.

4. ماکسیمم بدست آمده را با توجه به α و N با critical value مقایسه می‌کنیم و نتیجه‌ی تست را اعلام می‌کنیم.

تست chi square: در این تست اعداد را دسته‌بندی می‌کنیم و تعداد اعداد هر دسته را می‌شماریم (O_i) و با تعداد اعدادی که انتظار داشتیم در هر دسته باشد (E_i) مقایسه می‌کنیم:

$$X_0^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

$$E_i = np(x)$$

سپس عدد بدست آمده را از جدول، با توجه به مقدار α از سطر $n-1$ ام می‌خوانیم.

تست autocorrelation: یک مقدار i و یک مقدار m انتخاب می‌کنیم. از عدد i ام در لیست m تا m تا، تا انتهای جلو می‌رویم. سه مقدار زیر را محاسبه می‌کنیم:

$$\hat{\rho}_{im} = \frac{1}{M+1}\left[\sum_{k=0}^{M} R_{i+km}R_{i+(k+1)m}\right] - 0.25$$

$$\hat{\sigma}_{\rho_{im}} = \frac{\sqrt{13M+7}}{12(M+1)}$$

$$Z_0 = \frac{\hat{\rho}_{im}}{\hat{\sigma}_{\hat{\rho}_{im}}}$$

سپس درصد اطمینان مورد نیاز خود را از روی جدول پیدا می‌کنیم و با جمع سطر و ستون آن، بازه‌ی مجاز برای Z_0 را بدست می‌آوریم.

فصل هفت

در فصل قبل تولید اعداد با توزیع یکنواخت را آموختیم. در این فصل با استفاده از آن اعداد، مقادیر تصادفی با توزیع مشخص می‌سازیم.

روش inverse transform: بهتر است روی توزیع‌هایی استفاده شود که CDF وارون‌پذیر دارند. مراحل به صورت زیر هستند:

$$R_i \sim U[0,1] \Longrightarrow R_i = F_X(x) \Longrightarrow x = F_X^{-1}(R_i)$$

توابع مولد چند توزیع مختلف به شرح زیر هستند:

$$exponential\text{: } x = -\frac{1}{\lambda}\ln{(R)}$$

$$uniform(a,b)\text{: } x = a + (b-a)R$$

$$weibull\text{: } x = \alpha[-\ln(1-R)]^{\frac{1}{\beta}}$$

$$triangular\text{: } x = \begin{cases} \sqrt{2R}, & 0 \le R \le \frac{1}{2} \\ 2-\sqrt{2(1-R)}, & \frac{1}{2} < R \le 1 \end{cases}$$

$$geometric\text{: } X = q + \left\lceil \frac{\ln{(1-R)}}{\ln{(1-p)}} - 1 \right\rceil$$

برای توزیع‌های *empirical* دو راه داریم. در راه اول سعی می‌کنیم نمودار داده‌ها را رسم کنیم و از روی نمودار مقادیر جدید را استخراج کنیم.

در روش دوم، در صورت کوچک بودن اندازه‌ی مجموعه داده، *CDF* داده‌ها را رسم می‌کنیم (که شامل خطوط منقطع است) و سپس از فرمول زیر استفاده می‌کنیم:

$$\frac{j-1}{n} < R \le \frac{j}{n}$$

$$x = y_{j-1} + a_j\left[R - \frac{j-1}{n}\right]$$

که در آن a_j شیب خط j ام در نمودار *CDF* است:

$$a_j = \frac{y_j - y_{j-1}}{\frac{1}{n}}$$

به زبان ساده‌تر، مقدار R را روی نمودار عمودی مشخص می‌کنیم و در امتداد افقی ادامه می‌دهیم تا نمودار CDF را قطع کند. سپس نقطه‌ی متناظر x را به عنوان جواب انتخاب می‌کنیم.

در صورت زیاد بودن اندازه‌ی مجموعه داده، از همین روش استفاده می‌کنیم با این تفاوت که بازه‌ها را به صورت دستی تثبیت می‌کنیم (در روش قبلی خود اعداد بازه‌ها را تشکیل می‌دادند) و همچنین CDF را با استفاده از تعداد اعداد در هر بازه بدست می‌آوریم.

$$c_{j-1} < R < c_j$$

$$x = y_{j-1} + a_j\left[R - c_{j-1}\right]$$

$$a_j = \frac{y_j - y_{j-1}}{c_j - c_{j-1}}$$

نکته‌ی دیگری که حائز اهمیت است، نقطه‌ی شروع اولین بازه است که در این حالت برابر مینیمم عدد دیده شده و در حالت قبلی برابر ۰ می‌باشد.

روش *acceptance-rejection*: زمانی که فرم بسته‌ی CDF را نداشته باشیم از این روش استفاده می‌کنیم. ابتدا باید یک شرط پذیرش برای متغیرهای تصادفی بیابیم و سپس اعداد تصادفی تولید کنیم و ببینیم آیا در آن شرط قرار می‌گیرند یا خیر. شرط چند توزیع مختلف به صورت زیر است:

$$uniform\colon\ a \le R \le b$$

$$poisson\colon\ \prod_{i=1}^{n} R_i \ge e^{-\lambda} \ge \prod_{i=1}^{n+1} R_i$$

روش تقریبی برای تولید متغیرهای پوآسون به کمک توزیع نرمال:

$$\begin{aligned} z_1 &= (-2\ln R_1)^{\frac{1}{2}}\cos(2\pi R_2) \\ z_2 &= (-2\ln R_1)^{\frac{1}{2}}\sin(2\pi R_2) \\ \Longrightarrow N &= \left\lceil \lambda + \sqrt{\lambda} Z - 0.5 \right\rceil \end{aligned}$$

تکنیک *thining* برای *NSPP*:

1. $\lambda^* = \max\lambda(t)\,, t = 0, i = 1$

$$2.\ E \sim \exp(\lambda^*),\ t = t + E$$
$$3.\ R \sim U[0,1] \rightarrow R \leq \frac{\lambda(t)}{\lambda^*}? \Longrightarrow \tau_i = t, i = i + 1$$
$$4.\ return\ to\ 2.$$

تکنیک *convolution*: برای توزیع ارلانگ و دوجمله‌ای استفاده می‌شود.

$$X = \sum_{i=1}^{k} -\frac{1}{k\theta} \ln R_i$$

فصل هشت

پیدا کردن توزیع ورودی‌ها: از هیستوگرام استفاده می‌کنیم. پیشنهاد می‌شود طول بازه‌ها برابر ریشه‌ی تعداد داده‌ها باشد. سپس با استفاده از *Context* ورودی‌ها و شکل هیستوگرام، نوع توزیع را حدس می‌زنیم.

توزیع‌های مختلف و کاربردهایشان:

دو جمله‌ای	تعداد موفقیت‌ها در n آزمایش
دوجمله‌ای منفی و هندسی	تعداد آزمایش مورد نیاز برای مشاهده‌ی k موفقیت
پواسون	تعداد رویدادها در یک بازه‌ی زمانی / مکانی
نرمال	هر متغیری که مجموع چندین کامپوننت مختلف باشد.
لاگ-نرمال	هر متغیری که حاصل‌ضرب چندین کامپوننت مختلف باشد.
نمایی	مدت زمان بین دو رویداد
گاما	مدل کردن هر متغیر نامنفی
بتا	مدل کردن هر متغیر محدود به دو مقدار
ارلانگ	مجموع چند متغیر نمایی

ویبول	مدل کردن TTF
یکنواخت	احتمال هر رویداد یکسان است
مثلثی	حالتی که فقط مینیمم و ماکسیمم را می‌دانیم (حالت قوی‌تر یکنواخت)

نمودارهای *Qunatile-Quantile*: از داده‌ها نمونه‌برداری می‌کنیم. همچنین مقادیر $F^{-1}\left(\frac{j-0.5}{n}\right)$ از توزیعی که حدس می‌زنیم استخراج می‌کنیم. نمودار این دو را رسم می‌کنیم. اگر حدودا خط صاف بود، یعنی حدسمان درست بوده.

پارامترهای اطلاعات جمع‌آوری شده:

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}, \qquad S^2 = \frac{\sum_{i=1}^{n} X_i^2 - n\bar{X}^2}{n-1}$$

و اگر اطلاعات دسته‌بندی شده باشد:

$$\bar{X} = \frac{\sum_{j=1}^{c} f_j m_j^2}{n}, \qquad S^2 = \frac{\sum_{j=1}^{c} f_j m_j^2 - m\bar{X}^2}{n-1}$$

که در آن m_j میانگین دسته‌ی jام است.

اسلاید ۲۳ لکچر ۸ تخمین‌گرها مطالعه شود.

تخمین نرخ در *NSPP*:

$$\hat{\lambda}(t) = \frac{1}{n\Delta t}\sum_{j=1}^{n} C_{ij}$$

کوواریانس کوریلیشن دو متغیر:

$$cov(X_1, X_2) = E(X_1X_2) - \mu_1\mu_2$$

$$\rho = corr(X_1, X_2) = \frac{cov(X_1, X_2)}{\sigma_1\sigma_2}$$