Phuoc Tran-Gia and Tobias Hoßfeld

Performance Modeling and Analysis of Communication Networks

A Lecture Note

Version date: February 22, 2022
Online access: http://modeling.systems

2 Fundamentals and Prerequisites

Question and Answer

Question 2.1. A single server queue serves customers in random order. This scheduling discipline is called ROS (random order service). Are you allowed to apply Little's law?

a. Yes, Little can be applied to any system.

b. Yes, Little can also be applied for ROS.

c. Yes, it is a single server queue.

d. No, this is not possible.

Question 2.2. A system consists of two servers and a single queue. Again, incoming customers are served in random order. Are you allowed to apply Little's law?

a. Yes, Little can be applied to any system.

b. Yes, Little can also be applied for ROS.

c. No, it is not a single server queue.

d. No, this is not possible.

Question 2.3. A system consists of a single server and a single queue. The mean service time is $\mathrm{E}[\,B\,] = 10\,\mathrm{s}$. The arrival rate is $\lambda = 0.5\,\mathrm{s}^{-1}$. The mean number of customers in the system is $\mathrm{E}[\,X\,]$. Are you allowed to apply Little's law?

a. Yes, $\mathrm{E}[\,X\,] = \lambda \mathrm{E}[\,B\,]$.

b. Yes, if the mean sojourn time is known.

c. Yes, $\mathrm{E}[\,X\,] = \lambda(\mathrm{E}[\,B\,] + \mathrm{E}[\,W\,])$.

d. No, the system is not stable.

Question 2.4. Consider a GI/GI/n-∞ queue. The mean interarrival time $\mathrm{E}[\,A\,]$, the mean service time $\mathrm{E}[\,B\,]$ and the mean waiting time $\mathrm{E}[\,W\,]$ are known. Can you use Little's law to derive the mean number $\mathrm{E}[\,X_W\,]$ of customers waiting in the queue?

a. Yes, $\mathrm{E}[\,X_W\,] = \mathrm{E}[\,B\,]/\mathrm{E}[\,A\,]$.

b. Yes, $\mathrm{E}[\,X_W\,] = \mathrm{E}[\,W\,]/\mathrm{E}[\,A\,]$.

c. Yes, $\mathrm{E}[\,X_W\,] = (\mathrm{E}[\,B\,] + \mathrm{E}[\,W\,])/\mathrm{E}[\,A\,]$.

d. No, information is missing.

Question 2.5. Consider a GI/GI/n-0 loss system. The arrival rate λ and the mean service time $\mathrm{E}[\,B\,]$ are known. Can you use Little's law to derive the mean number $\mathrm{E}[\,X\,]$ of customers?

a. Yes, $\mathrm{E}[\,X\,] = \lambda \mathrm{E}[\,B\,]$.

b. Yes, $\mathrm{E}[\,X\,] = \lambda \mathrm{E}[\,B\,]/n$.

c. Yes, $\mathrm{E}[\,X\,] = \lambda \mathrm{E}[\,B\,] \cdot n$.

d. No, information is missing.

Exercises and Problems

Random Variables and Distributions

Problem 2.1. Dice A symmetrical dice is rolled. The random variable X has the value '0' for an odd number and the value '1' for an even number. The random variable Y is '0' for a low number (1,2 or 3) and '1' for a high number (4,5 or 6). What are the distributions of X and Y and the joint distribution for (X, Y)? Are X and Y positively or negatively correlated?

Problem 2.2. Hard Disk Drive An infrastructure-as-a-service provider implements a data storage system using hard disk drives (HDD) with the following parameters:

- 2 platters with 615 tracks where each platter has two read/write heads, one for the top of the platter and one for the bottom,
- $3\,\mathrm{ms}$ positioning time from track to track,
- 34 sectors per track,
- rotation speed of 3600 revolutions per minute (RPM).

The switching time between the four read/write heads is negligible.

2.2.1. Assuming that the track is already set correctly, the access time Z is analyzed. The access time is the time span from the access request until the beginning of the sector is reached. What is the probability density function $z(t)$ and the cumulative distribution function $Z(t)$ of the random variable Z? What are the expected value, the variance, the standard deviation and the coefficient of variation of Z?

2.2.2. Furthermore, the time span A is considered which is required for driving into the correct track. The expected value and the variance of the time span A is to be determined. Two cases should be considered: the read/write head is positioned above track 0 and track 307, respectively. What is the expected value of A for the general case, where it is assumed that all tracks as start and destination tracks are equally likely?

☞ *Hint:* $\sum\limits_{i=0}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$

Problem 2.3. Binomial and Poisson Distribution Show that for large N and for small values of $p = \mu/N$ the binomial distribution $X \sim \mathrm{BIN}(N, p)$ is approximated by the Poisson distribution $Y \sim \mathrm{POIS}(\mu)$. The expected values are identical and it is $\mathrm{E}[\,X\,] = \mathrm{E}[\,Y\,] = \mu$.

Problem 2.4. Comparions of RVs A<B Let two non-negative, independent random variables A and B be given, which are exponentially distributed with the rate λ and μ, respectively. Derive the following relationships:

$$P(A < B) = \frac{\lambda}{\lambda + \mu}.$$

☞ *Hint:* The integral of the two-dimensional joint probability density function is to be derived for the range $A < B$.

Problem 2.5. Conditional Random Variable Three statistically independent, continuous random numbers A_1, A_2 and A_3 are considered, which are exponentially distributed with the mean values $\frac{1}{\lambda_1}$, $\frac{1}{\lambda_2}$ and $\frac{1}{\lambda_3}$.

2.5.1. Derive the probability that $A_1 \leq A_2 \leq A_3$ based on λ_1, λ_2 and λ_3.

2.5.2. What is the PDF of the conditional random variable $A = A_2|(A_1 \leq A_2 \leq A_3)$? What is the expected value $\mathrm{E}[\,A\,]$ and the variance $\mathrm{VAR}[\,A\,]$?

Problem 2.6. Birthday Problem We consider the so-called birthday problem which asks for the probability that, in a set of randomly chosen people, at least two will share a birthday. In this exercise, derive the following probabilities.

2.6.1. There are n people in a lecture hall with 50 people who have birthdays today ($n = 0, 1, 2$).

2.6.2. At least two people in a room with 50 people have birthdays today.

February 29th is not taken into account. Thus, a year consists of 365 days and all birthdays are assumed to be equally likely. ☞ *Hint:* For large n, Stirling's formula applies: $n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$

Problem 2.7. Faulty Machine Processing A machine processes individual workpieces in exponentially distributed time. The mean service time is μ^{-1}. The service times are independently and identically distributed (iid). A quality control is carried out after each operation. With the probability q, a workpiece is faulty and must be processed again by that machine. Accordingly, a workpiece successfully passes the quality control with probability $p = 1 - q$ and gets to the next machine.

2.7.1. What is the distribution of the number X of machine operations on a single workpiece?

2.7.2. Which distribution does the *total* service time Y follow? What is the mean value $\mathrm{E}[\,Y\,]$ of the total service time Y that a single workpiece experiences?

Problem 2.8. Mutually Exclusive Event System Let $(B_i)_{0 \leq i < n}$ be a mutually exclusive event system and A a random variable. Show or refute the following statements and quote the rules of probability theory necessary for the transformations.

2.8.1. $E[A] = \sum_{0 \leq i < n} P(B_i) E[A|B_i]$

2.8.2. $E[A^2] = \sum_{0 \leq i < n} P(B_i) E[(A|B_i)^2]$

2.8.3. $VAR[A] = \sum_{0 \leq i < n} P(B_i) VAR[A|B_i]$

Problem 2.9. Erlang-k Distribution Show that for an Erlang-k distributed random variable A with parameters k and λ the following equations are valid:

2.9.1. $E[A] = \frac{k}{\lambda}$

2.9.2. $c_A = \frac{1}{\sqrt{k}}$

Problem 2.10. Uniform Distribution The random variables X and Y are statistically independent and identically distributed in the range $[a_x; b_x] = [3; 5]$ and $[a_y; b_y] = [2; 7]$, respectively. Draw the corresponding probability density functions. Determine the maximum range R_0 or R_1, in which the joint cumulative distribution function $Z(x, y) = P(X \leq x, Y \leq y)$ takes the value 0 or the value 1. What is the probability $P(R_0)$ and $P(R_1)$, respectively? What is the joint probability density function?

Problem 2.11. Fair and Unfair Coin Given are two coins. One coin is fair; the probability of *head* as the result of a throw is $1/2$. The other coin is unfair; the probability for *head* is $p > 1/2$. After one of the coins has been selected at random, this coin is tossed n times and *head* will appear n times.

2.11.1. Given this result, what is the probability that the selected coin is the unfair coin?

2.11.2. Based on the outcome of the above experiment, it is decided to identify the chosen coin as the unfair coin. How big would n have to be so that the probability of being wrong with this decision is at most 0.001?

Problem 2.12. Slot-based Communication Protocol We consider a slot-based communication protocol in which packets can only arrive at the beginning of a time slot. The slot length is given by Δt. At every possible arrival time $k \cdot \Delta t$ there is an arrival with probability p, with probability $1 - p$ no packet will arrive.

2.12.1. What is the distribution A for the interarrival time of the packets? Determine the mean value $\mathrm{E}[\,A\,]$ and the coefficient of variation c_A. What is the number Y of packets in an interval of n slots? What is $\mathrm{E}[\,Y\,]$ and c_Y?

2.12.2. Now assume that more than one packet can arrive in a slot. With probability p an arrival will take place at the beginning of a time slot; with probability p, the next arrival will be in the same time slot. What is the distribution A for the interarrival time of the packets? Determine the mean value $\mathrm{E}[\,A\,]$ and the coefficient of variation c_A. What is the number K of packets which arrive in the same time slot?

2.12.3. Assume again that several arrivals can occur at the beginning of a time slot. However, if there is more than one arrival within a time slot, a collision occurs and none of the packets is correctly received. What is the distribution for the interarrival time of correctly received packets? What is the number of correctly received packets in an interval of n slots? What do you observe when comparing the mean values in this subproblem with the mean values in subproblems 2.12.1. and 2.12.2..

Problem 2.13. Packet-switched and Circuit-switched Network Two hosts A and B are communicating in a packet-switched communication network, as visualized in the figure below. There are three different routing paths R_1, R_2, R_3 which are randomly selected with probability $p_1 = 0.3, p_2 = 0.4, p_3 = 0.3$.

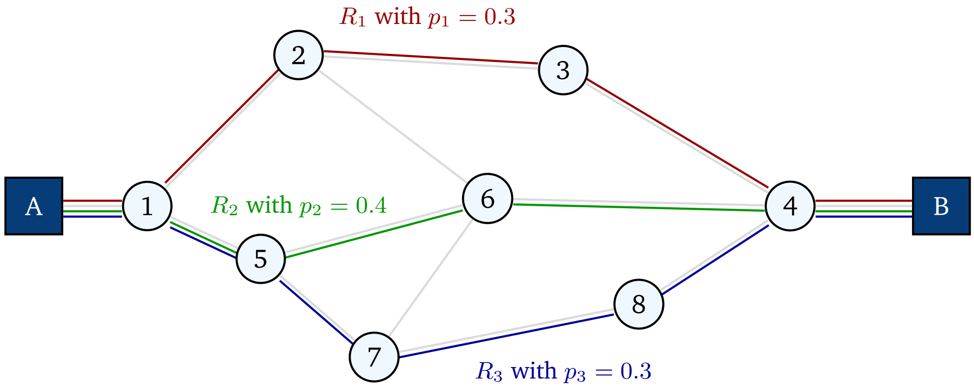

Figure 2.1: Packet switched network with three different routes R_1, R_2, R_3.

The transmission time for any packet on the path R_i from A to B is described by the continuous random variable T_i and the CDF $F_i = P(T_i \leq t)$. We assume that the transmission times are independent of each other.

$$F_1(t) = P(T_1 \leq t) = 1 - e^{-t/t_1} \qquad \text{with} \quad t_1 = 2\,\text{ms}$$
$$F_2(t) = P(T_2 \leq t) = 1 - e^{-t/t_2} \qquad \text{with} \quad t_2 = 1\,\text{ms}$$
$$F_3(t) = P(T_3 \leq t) = \begin{cases} 0, & t < 5\,\text{ms} \\ 1, & t \geq 5\,\text{ms} \end{cases}$$

2.13.1. Determine the mean value $\mathrm{E}[\,T\,]$, the variance $\mathrm{VAR}[\,T\,]$ and the coefficient of variation c_T for the the transmission time T between A and B.

2.13.2. A message on application layer requires the transmission of two data packets from A to B. Determine the mean value $\mathrm{E}[\,T_m\,]$ of the total transmission time T_m of a single message.

2.13.3. The host A sends the two packets P_1 and P_2 of a message simultaneously over two different routes. P_1 uses the route R_1 and P_2 uses the route R_2, respectively. Calculate the probability that the packet P_1 is overtaken by P_2.

☞ *Hint:* From the law of total probability it follows for an event A and the random variable T with the distribution density function $f(t)$:

$$P(A) = \int_{t=0}^{\infty} P(A|T = t) f(t)\; dt\;.$$

2.13.4. Instead of a packet-switched network, we assume now that dedicated virtual connections are used to avoid overtaking of the two packets of a single message. What is the expected value $\mathrm{E}[\,T_v\,]$ of the total transmission time of a message, if the virtual connection is using the routes R_i with the corresponding probability p_i for $i = 1, 2, 3$. What is the relation between $\mathrm{E}[\,T_v\,]$ and $\mathrm{E}[\,T_m\,]$.

Problem 2.14. P2P File Sharing In P2P file sharing networks such as BitTorrent or eDonkey, larger files are divided into chunks. To download a file, these chunks can be downloaded from several peers at the same time. As soon as a peer has completely downloaded a chunk, it becomes the source itself and offers the chunk to other peers for download.

We consider a file consisting of N chunks. We assume that a peer A starts the download of the file and downloads different chunks from N different peers in parallel. The transmission time for chunk i is described by a random variable C_i. Assume that C_i is exponentially distributed with the parameter λ:

$$c_i(t) = \frac{d}{dt}P(C_i \leq t) = \lambda e^{-\lambda t}\ , \text{ for } t \geq 0.$$

2.14.1. The random variable S describes the duration until peer A itself becomes the source for one of the N chunks. Assuming the statistical independence of (C_i), calculate the probability density function $\frac{d}{dt}S(t)$. What are the mean $\mathrm{E}[\,S\,]$ and the variance $\mathrm{VAR}[\,S\,]$? Which distribution describes S and what are its parameters?

2.14.2. Assume that $N = 2$ and $\lambda = 1\,\mathrm{min}$. The random variable D describes the total download time. What is the cumulative distribution function $D(t)$? What is the mean $\mathrm{E}[\,D\,]$?

Problem 2.15. VoIP An Internet telephony software uses P2P technology to enable calls between a peer A and a peer B, which are each located behind a firewall. Then the packets between A and B are relayed via a third peer C, which is not behind a firewall.

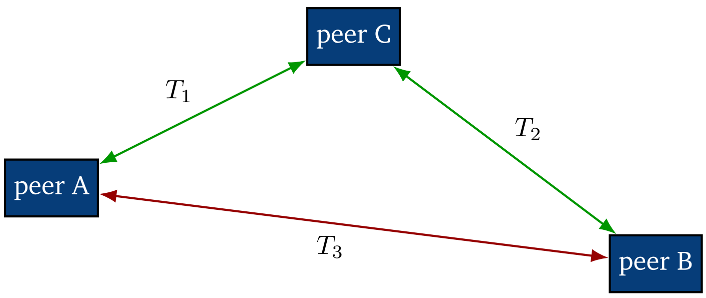

Figure 2.2: Transmission time of packets in a peer-to-peer network.

Assume that the packet transmission time T_1 (in ms) from peer A to peer C and T_2 from peer C to peer B follows a shifted geometric distribution: $T_i \sim \mathrm{GEOM}_m(p)$ with $m = 10$ and $p = 0.1$ for $i = 1, 2$. Furthermore, T_1 and T_2 are statistically independent. The shifted geometric distribution GEOM (m, p) is defined as follows:

$$x(i) = \begin{cases} (1-p)^{i-m} \cdot p & \text{for } i = m, m+1, \ldots, \\ 0 & \text{otherwise.} \end{cases}$$

What is the distribution $t(i) = P(T = i)$ of the total packet transmission time T from peer A to peer B via the relay peer C? Calculate the mean $\mathrm{E}[\,T\,]$ and the variance $\mathrm{VAR}[\,T\,]$. Which distribution does $T' = T - 2m$ follow and what are their parameters?

Problem 2.16. Email Server A mail client contacts the server periodically at time intervals of $\Delta t = 5$ min. With probability p there is a new email in the inbox; with probability $q = 1 - p$ no new email has been received. For each of the following subproblems, answer the following questions. What is the distribution of the random variable under consideration? How is the distribution called? What is the expected value of the random variable?

2.16.1. The random variable Y describes all possible events at any time interval.

2.16.2. The random variable X describes the number of new emails within one day.

2.16.3. The random variable A describes the interarrival time of new messages.

2.16.4. Exactly every second message is a spam mail. The random variable B describes the interarrival time of incoming non-spam.

Transform Methods

Problem 2.17. Laplace-Stieltjes Transform For the parametric investigation of traffic models, distributions of random variables are often required for which only two parameters are known (mean value and coefficient of variation). For different coefficients of variation c, consider the following classes of continuous distributions:

(1) deterministic distribution $A_1 \sim \mathrm{D}(t_1)$ with CDF $A_1(t) = \begin{cases} 0\,, & t < t_1 \\ 1\,, & t \geq t_1 \end{cases}$;

(2) exponential distribution $A_2 \sim \mathrm{EXP}(\lambda_2)$ with parameter λ_2;

(3) hyperexponential distribution of second order A_3 with parameters $\lambda_{3,i}$ $(i = 1, 2)$ and symmetry assumption;

(4) sum of two independent phases $A_4 = A_{4,1} + A_{4,2}$, a deterministic phase (D) and an exponential phase (M): $A_{4,1} \sim \mathrm{D}(t_4)$ and $A_{4,2} \sim \mathrm{EXP}(\lambda_4)$

2.17.1. Solve the task now for the four distributions $A_i,\ i = 1, 2, 3, 4$. What is the probability density function $a_i(t)$? Calculate the Laplace-Stieltjes transform $\Phi_i(s)$ of the r.v. A_i. Derive the mean values $\mathrm{E}[\,A_i\,]$ and coefficients of variation c_{A_i} from $\Phi_i(s)$.

2.17.2. Measurements of the transmission time T (r.v.) of data packets in a LAN have revealed the following behavior:

$$T = \begin{cases} 1\,\mathrm{ms} & \text{with probability } 0.8\,, \\ 10\,\mathrm{ms} & \text{with probability } 0.2\,. \end{cases}$$

Calculate the parameters of the four distributions above in such a way that $\mathrm{E}[\,T\,]$ and c_T are approximated as well as possible. Which of the approximating distributions (1) to (4) appears most suitable?

2.17.3. Draw the cumulative distribution functions of $A_i(t)$ and $T(t)$.

Problem 2.18. Compound Distribution A compound distribution is a probability distribution that results from assuming that a random variable is distributed according to a parametrized distribution where the parameter of that distribution is also a random variable.

Given a discrete random variable X with the distribution $x(i) = P(X = i)$, $i \in \mathbb{N}_0$, and the generating function $X_{GF}(z) = \sum_{i=0}^{\infty} x(i) \cdot z^i$. A new random variable Y is now formed as the sum of N consecutive random variables $X_1, X_2, \ldots, X_N$: $Y = X_1 + X_2 + \ldots + X_N$ and $X_i \sim X$ for $i = 1, 2, \ldots, N$. The number N of summands is itself a discrete random variable with distribution $n(i) = P(N = i)$, $i \in \mathbb{N}_0$, and the generating function $N_{GF}(z) = \sum_{i=0}^{\infty} n(i) \cdot z^i$. Then, $y(i) = P(Y = i)$ is called *compound probability distribution*; $x(i)$ is referred to as the *inner distribution* and N as the *outer distribution*.

2.18.1. Show that the generating function $Y_{GF}(z) = \sum_{i=0}^{\infty} y(i) \cdot z^i$ of the random variable Y holds:

$$Y_{GF}(z) = N_{GF}(X_{GF}(z)) \ . \tag{2.1}$$

☞ *Hint:* First consider the conditional distribution $P(Y = j|N = i)$ with i kept constant and then apply the law of total probability.

2.18.2. Show with the help of $Y_{GF}(z)$ from Equation (2.1) that the mean value and the variance satisfy the following equations:

$$\mathrm{E}[\,Y\,] = \mathrm{E}[\,N\,] \cdot \mathrm{E}[\,X\,] \ , \tag{2.2}$$

$$\mathrm{VAR}[\,Y\,] = \mathrm{E}[\,N\,] \cdot \mathrm{VAR}[\,X\,] + \mathrm{VAR}[\,N\,] \cdot \mathrm{E}[\,X\,]^2 \ . \tag{2.3}$$

2.18.3. Instead of the discrete random variable X, consider a continuous random variable T with the Laplace-Stieltjes transformation (LST) $\Phi_T(s)$. The new random variable Y is formed according to $Y = T_1 + T_2 + \ldots + T_N$, where all T_i, $i = 1, 2, \ldots, N$, follow the same distribution T and have the same LST $\Phi_T(s)$. Show that for the LST $\Phi_Y(s)$, the mean value $\mathrm{E}[\,Y\,]$ and the variance $\mathrm{VAR}[\,Y\,]$ of the random variable Y the following equations apply:

$$\Phi_Y(s) = N_{GF}(\Phi_T(s)) \ , \tag{2.4}$$

$$\mathrm{E}[\,Y\,] = \mathrm{E}[\,N\,] \cdot \mathrm{E}[\,T\,] \tag{2.5}$$

$$\mathrm{VAR}[\,Y\,] = \mathrm{E}[\,N\,] \cdot \mathrm{VAR}[\,T\,] + \mathrm{VAR}[\,N\,] \cdot \mathrm{E}[\,T\,]^2 \ . \tag{2.6}$$

Problem 2.19. Properties of the Laplace Transform Prove the following two properties of the Laplace transform:

2.19.1. Laplace transform of CDF: $\mathrm{LT}(A(t)) = \frac{1}{s}\Phi_A(s), s \neq 0$

2.19.2. First derivative property of the Laplace transform: $\mathrm{LT}\big(\frac{d}{dt}a(t)\big) = s\Phi_A(s) - a(0)$

☞ *Hint:* Use the properties of the integral and well-known integral arithmetic tricks.

Problem 2.20. Error-prone Transmission Channel Messages of random length are to be transmitted in packets with a constant packet size of s bits. The transmission channel is error-prone and is characterized by the independent bit error probability p_b. The number N of packets that belong to a message follows a Poisson distribution with an average of M blocks per message. The number of corrupted packets per message is to be analyzed. We assume that a single bit error results in a packet error.

2.20.1. What is the packet error probability p_0?

2.20.2. The disruption or non-disruption of a packet during transmission is identified by the two-valued (Bernoulli) random variable X. Give the distribution $x(i)$ of the random variable X and its probability generating function $X_{GF}(z)$.

2.20.3. What is the distribution $n(i)$ and the probability generating function $N_{GF}(z)$ of the random variable N, which describes the number of packets per message?

2.20.4. What is the generating function $Y_{GF}(z)$ of the number Y of corrupted packets per message? ☞ *Hint:* See Problem 2.18.

2.20.5. What is the mean $\mathrm{E}[\,X\,]$, the variance $\mathrm{VAR}[\,Y\,]$ and the distribution $y(i)$ of the r.v. Y?

☞ *Hint:* Reshape $Y_{GF}(z)$ so that it is of the form $\sum\limits_{i=0}^{\infty} f(i) \cdot z^i$ and read $y(i)$ directly.

Problem 2.21. Probability Generating Function

2.21.1. X is the sum of two statistically independent discrete random variables X_1 and X_2 with the probability distributions $x_1(i)$ and $x_2(i)$. Show with the help of the Cauchy product formula that for the generating function of the distribution $x(i) = P(x = i)$ the *convolution theorem* applies:

$$X_{GF}(z) = X_{1,GF}(z) \cdot X_{2,GF}(z) \; .$$

2.21.2. Define the distribution $y(i)$ of the random variable Y as follows: $y(i) = x(i - k)$, where $x(i) = 0$ for $i < 0$. Show the so-called *displacement theorem*:

$$Y_{GF}(z) = z^k \cdot X_{GF}(z).$$

☞ *Hint:* Cauchy product formula: $\left(\sum\limits_{n=0}^{\infty} a_n\right) \cdot \left(\sum\limits_{n=0}^{\infty} b_n\right) = \sum\limits_{n=0}^{\infty} \sum\limits_{k=0}^{n} a_k b_{n-k}$

Problem 2.22. Forwarding of Messages Messages arrive at a network node K of a communication system. The interarrival times A of messages are independent of each other and exponentially distributed with the parameter λ. From this network node K, n further network nodes can be reached, including node A. With probability p these messages are forwarded to the node A, with probability $1-p$ to any of the other nodes.

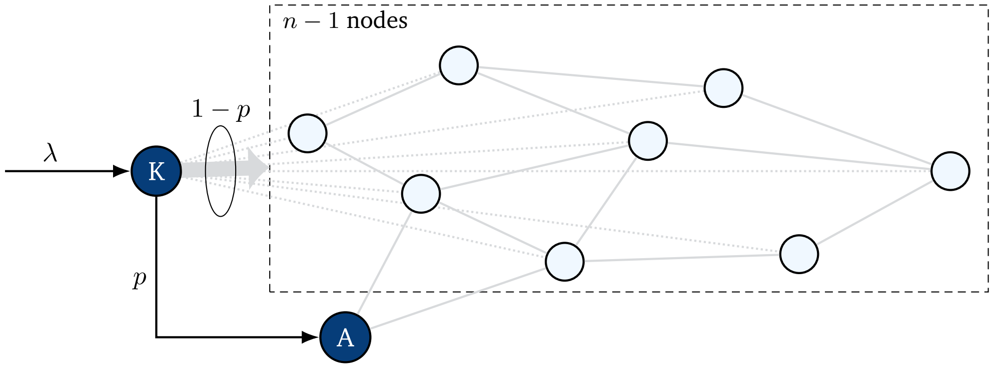

Figure 2.3: Node K sends a message to node A with probability p and to one of the other $n-1$ nodes with probability $1-p$.

The random variable X specifies the time between successive messages that leave the network node K with destination node A. This corresponds to the time from the arrival of a message with destination A and the arrival of the immediately next message with destination A. The processing time of network node K is neglected.

2.22.1. What is the mean $\mathrm{E}[X]$ and the coefficient of variation c_X? What distribution does the random variable X follow?

2.22.2. Now assume that the forwarding is not random, but rather deterministically based on the round robin principle for load distribution reasons. Hence, the packet j is forwarded to node $i = j \bmod n$. Determine the mean, coefficient of variation and the distribution of X.

Problem 2.23. Cox Distribution Similar to the Erlang-k distribution, the Cox-k distribution is based on k consecutive exponentially distributed phases of duration A_j with parameter λ_j for $j = 1, \ldots, k$. The phase process is always started in phase 1 and the process stays there for time A_1. After the end of the time in phase j, the process stops with a given probability p_j for $1 \leq j < k$ without entering the next phase $j+1$. It is $p_k = 1$.

2.23.1. Sketch the phase representation of the Cox-k distribution with parameters λ_j and p_j.

2.23.2. What is the Laplace transform of the PDF of the Cox-k distribution?