Simulation and Modeling I

Assignment 2

Problem 1 (1 + 0.5 + 1 points):

In unsorted sets with n different items, we perform sequential and random searches for one particular item (of which we know that it is included exactly once in each set).
Sequential search means that we pick one item after the other from the set until the desired item is found.
Random search means that we check an arbitrary item from the set, but leave it in the set, and recheck until the desired item is found.

a) What is the **mean number of checks** (including accidental rechecks in random search) until the item is found in a set **in either case** (with sequential and with random search)?

b) Which **distribution** does the random variable "number of checks" have **in each case**? Simply give the names and characteristic parameters of these distributions.

c) For the distribution arising from random search: show **formally** that this distribution enjoys the **memoryless property** (in analogy to the bottom of slide 33 of the exercise video "Basics of Probability Theory")! Interpret this memoryless property for random search **informally** (i.e., in your own words).

Problem 2 (1 + 2.5 points):

In technical systems and software programs, processes are often started at the same time and run in parallel. Before the next phase of the system/program can be entered, sometimes all of these processes have to be finished or sometimes simply a single one. We define the runtime of process **k** with the random variable X_k.

Let $X_1, X_2, \ldots, X_n$ be (mutually) independent and let $Y = \mathbf{min}\{X_1, X_2, \ldots, X_n\}$ and $Z = \mathbf{max}\{X_1, X_2, \ldots, X_n\}$. What do Y and Z model?

a) Show that the distribution functions $\mathbf{F_Y}$ of Y and $\mathbf{F_Z}$ of Z are given by
$F_Y(t) = 1 - \prod_{i=1}^{n}(1 - F_{X_i}(t))$ and $F_Z(t) = \prod_{i=1}^{n} F_{X_i}(t)$, respectively.

Hints: You may find it easier to show the relationship for Z first...
For independent random variables A and B, we have:
$P(A \leq t \text{ and } B \leq t') = P(A \leq t) \cdot P(B \leq t')$

b) Specialize (and simplify) these distribution functions for **independent and identically exponentially distributed X_i** (i=1,...,n)!
What distribution do you get for **Y**? What is the value of **E[Y]** in this case (exponentially distributed X_i)? Describe this observation in your own words.

Problem 3 (1.5 + 1 + 1 points):

A systems runs by a generator that provides power for the system. If the generator fails, the system has a battery, which can supply it with power for **exactly five more days**. Let the time to failure X for the generator be exponentially distributed with **expectation 1700** days, while Y denotes the time to the complete failure of the system.

a) What is the **probability** that the generator fails **within the first 1700 days**? What is the **quantile** with respect to this probability value **for random variable Y**?

b) Compute the **coefficient of variation for both X and Y**! Hint: X is exponentially distributed but Y is not represented by the "standard" exponential distribution. Think about how the distribution of Y would look like.

c) What is the (coefficient of) **correlation between X and Y**?

Problem 4 (1.5 + 1 points):

Assume a single-server queue (see lecture notes to *Introduction to Simulation*, 18ff) with one service unit (one packet can be processed at each time). For this single-server queue, the interarrival times of packets are exponentially distributed (with rate λ) and the service time for serving one packet is exponentially distributed (with rate μ). This is also called M/M/1 queue. The **mean number of customers** in this system (**including the service unit**) is computed as:

$$N = \frac{U}{1-U}, \text{ where } U = \frac{\lambda}{\mu}.$$

Besides, if we assume that the service time is generally distributed (arbitrary / any distribution), then the queue is called M/G/1 queues. We assume that the service times have the same mean as above ($1/\mu$). Then the **mean waiting time** (i.e., mean delay in the queue **excluding the service unit**) is given by

$$W = \frac{U(c_s^2 + 1)}{2\mu(1-U)},$$

where c_s is the coefficient of variation of the service time distribution.

a) Under which condition (on the arrival rate λ) do the mean delay (**D**) in the system (including the service unit) and the mean number (**N**) of customers in the system have the same numeric values for **an M/M/1 queue**? Hint: The mean delay in the system and the mean waiting time differ by the mean service time.

b) For identical mean service times, by which factor do the mean waiting times $\mathbf{W^{M/M/1}}$ and $\mathbf{W^{M/D/1}}$ differ for an M/M/1 and an M/D/1 queue? (D stands for deterministic / constant service time).

Just FYI: The formula

$$W = \frac{U(c_s^2 + 1)}{2\mu(1-U)}$$

is the <u>famous</u> Pollaczek-Khinchin mean value formula (PK formula). ☺