Tutorial 3 Solutions

1. Since the system of hard disks is set up in a RAID 0 configuration, the first error that occurs in any of the disks will break the disk array. We know that the error process for each of the K hard disks can be modeled as an (independent) Poisson process:

$$N_i(t) \sim \text{Poisson}(\lambda\, t)$$

   Consider now the superposition $Z(t)$ of these to form the error process of the complete disk array and let X be the random variable for the interarrival time for errors of the complete disk array. We have that

$$Z(t) = \sum_{i=1}^{K} N(t) \sim \text{Poisson}(K\, \lambda\, t),$$

   and therefore $X \sim \text{Exp}(K\, \lambda)$, from which follows the result

$$E[X] = \frac{1}{K\, \lambda} \quad *$$

   Note that here we have

$$\lambda = \frac{1}{3 * 10^5\text{h}} \approx \frac{1}{34.224\text{y}}.$$

   We obtain the following table of results (all values are in years).

Disk capacity (TB)	# Disks(K)	MTTF = $E[X]$ (Years)
16	4	8.562
64	16	2.140
256	64	0.535
1024	256	0.134
4096	1024	0.033

2. Since the system of hard disks is set up in a RAID 5 configuration (and broken disks are not replaced), the second error that occurs in any of the disks will break the disk array. As before, we know that the error process for each of the K hard disks can be modeled as an (independent) Poisson process. Consider now the time that has passed until the first failure in any of the disks (which is itself a random variable) and denote it by X_1. Denote by X_2 the time that passes between the first and the second failure (again, a random variable). Because of the linearity of the expected value we have

$$E[X] = E[X_1] + E[X_2],$$

   where X is the time until the complete system fails. Note that as before $E[X_1] = 1/(K\, \lambda)$. Consider now the time at which the first disk fails. Because of the memorylessness property of each of the $K-1$ remaining arrival interarrival processes, we know that their remaining interarrival time is still exponentially distributed.† Hence, we can again form the superposition process and obtain $E[X_2] = 1/((K-1)\, \lambda)$, so that

$$E[X] = \frac{1}{K\, \lambda} + \frac{1}{(K-1)\, \lambda} = \frac{2K-1}{K\, (K-1)\, \lambda}$$

   follows. We obtain the following table of results (note: we now need one more disk for storing the parity data).

Disk capacity (TB)	# Disks(K)	MTTF = $E[X]$ (Years)
16	5	12.557
64	17	3.917
256	65	1.046

pairs = K / 2;

for it $\leftarrow 1$ **to** repetitions **do**
minval $\leftarrow \infty$;

for $p \leftarrow 1$ **to** pairs **do** **Tmp** $\leftarrow$ Exponential(λ) // **Tmp** is a 2×1 array
maxval $\leftarrow$ Max(**Tmp**);
if minval > maxval **then** minval $\leftarrow$ maxval;
end if
end for

samples[it] $\leftarrow$ minval; **end for**

return Mean (samples);

Algorithm 1: Monte-Carlo type simulation for computing the mean value of the time to failure as defined by Equation (**??**). Note that the pseudocode repeatedly draws values from the exponential distribution with parameter λ.

It is important to note that

$$E[\min_{j=0}^{(K/2)-1} \max_{i \in \{2j,2j+1\}} X_i] \neq \min_{j=0}^{(K/2)-1} E[\max_{i \in \{2j,2j+1\}} X_i]$$

in general, so that equality only holds in exceptional cases (for example $K = 2$).

In the following we discuss two approaches for computing the expected value. The first approach, as indicated on the problem sheet, is by simulation. Algorithm **??** is a pseudocode formulation for a Monte-Carlo type simulation for computing the mean value for the time to failure based on repeated sampling for the individual failure times. Essentially, the algorithm generates samples for the time to failure for each of the mirror pairs by sampling two failure times for each disk inside each pair from the exponential distribution. These values are then aggregated according to the min-max evaluation outlined above. When running 500 repetitions we obtain the following results (all values are in years, as before).

Disk capacity (TB)	# Disks(K)	Median	Mean (MTTF (Years))	Standard Deviation
4	4	25.526	30.725	21.859
16	16	11.624	12.946	7.845
64	64	5.559	5.976	3.318
256	256	2.556	2.778	1.520
1024	1024	1.279	1.349	0.693

The second analytical approach (initially based on a suggestion by Jozsef Hegedus) aims at computing the expected value directly. This is complicated due to the fact that the maximum of two exponentially distributed random variables is not itself distributed according to an exponential distribution. However, starting from the cumulative distribution function for each of the disks, we are able to formulate the cumulative distribution function of the maximum lifetime of any of the two disks in any mirror pair. In summary, the equations are as follows.

$$\mathrm{Prob}(X_i \leq t) = 1 - e^{-\lambda t} \tag{2}$$

$$\begin{aligned}\mathrm{Prob}(\max\{X_i, X_{i+1}\} \geq t) &= 1 - \mathrm{Prob}(\min\{X_i, X_{i+1}\} \leq t) \\ &= 1 - (\mathrm{Prob}(X_i \leq t) * \mathrm{Prob}(X_{i+1} \leq t)) = 1 - (1 - e^{-\lambda t})^2\end{aligned} \tag{3}$$

$$\mathrm{Prob}(\min_j \max\{X_{2j}, X_{2j+1}\} \geq t) = (\mathrm{Prob}(\max\{X_{2j}, X_{2j+1}\} \geq t))^{\frac{K}{2}} = (1 - (1 - e^{-\lambda t})^2)^{\frac{K}{2}} \tag{4}$$

$$\mathrm{pdf}(t) = \frac{d}{dt}[1 - \mathrm{Prob}(\min_j \max\{X_{2j}, X_{2j+1}\} \geq t)] \tag{5}$$

$$E[X] = \int_0^\infty t \cdot \mathrm{pdf}(t)\ \mathrm{dt} \tag{6}$$

In our trials Maple was unable to compute a closed form expression for the integral in (**??**). However, we are able to obtain (approximate) values for some small fixed values for K and for the given value of lambda. The computed values are surprisingly close to those that were obtained by simulation. Note that in practice, particularly when systems are more complex than described here, simulation may be the only viable option.

Disk capacity (TB)	# Disks(K)	Mean (analytical)	Mean (experimental)
4	4	31.372	30.725
16	16	13.031	12.946
64	64	5.917	5.976
256	256	2.817	2.778