Marius Hofert, Riley Metzger University of Waterloo 2015-01-09

STAT 340: Computer Simulation of Complex Systems

Assignment 1 (due on 2015-01-23)

Exercise 1 (Distribution of the minimum of n independent random variables) (2 points)
Suppose that we have random variables $X_1, X_2, \dots, X_n$ which are mutually independent and have distribution functions $F_1, F_2, \dots, F_n$, respectively. Show that the distribution of the *minimum* $X_{(1)} = \min\{X_1, X_2, \dots, X_n\}$ is given by

$$F(x) = 1 - \prod_{i=1}^{n}(1 - F_i(x)).$$

Exercise 2 (Cauchy distribution) (2+2+2=6 points)
Let $f(x) = c/(1 + x^2)$, $x \in \mathbb{R}$, for some $c \in \mathbb{R}$.

a) Determine all c such that f is a density.

b) Compute the corresponding distribution function F.

c) Compute, if it exists, the expectation and the variance of F.
Hint. You do not have to, but you may use *Jensen's inequality*: For any random variable X and convex function φ, $\varphi(\mathbb{E}[X]) \le \mathbb{E}[\varphi(X)]$.

Exercise 3 (Almost sure equality vs equality in distribution) (4 points)
Let X and Y be two random variables. Show that $X = Y$ a.s. (i.e., $\mathbb{P}(X = Y) = 1$) implies that X and Y have the same distribution function. Does the converse also hold?

Exercise 4 (Gamma distribution) (2+3+2=7 points)
Let

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x), \quad x \ge 0,\ \alpha > 0,\ \beta > 0,$$

where $\Gamma(z) = \int_0^\infty x^{z-1} \exp(-x)\, dx$ denotes the *Gamma function*.

a) Show that f is a density on $[0, \infty)$.
Note. f is the density of the *Gamma distribution* with *shape* α and *rate* β, denoted by $\Gamma(\alpha, \beta)$.

b) Let $X \sim \Gamma(\alpha, \beta)$. Compute $\mathbb{E}[X^k]$, $k \in \mathbb{N}$, and $\mathrm{Var}[X]$.

c) Let $X \sim \Gamma(\alpha, 1)$. Show that $Y = X/\beta \sim \Gamma(\alpha, \beta)$.

Exercise 5 (Random variables are constant if and only if they have 0 variance) (3 points)
Let X be a random variable with $\mathbb{E}[X^2] < \infty$. Show that $\mathrm{Var}\, X = 0$ if and only if $X = \mathbb{E}X$.

Exercise 6 (Birthday problem) (2+3=5 points)
Suppose there are k persons in a room.

a) Find the probability that no two persons in the room have the same birthday.

b) Compute in R, for each reasonable k, the probability that at least two people have the same birthday and plot your result as a function of k. Determine the smallest k for which it is more than likely (i.e., with probability greater than 0.5) that at least two people share the same birthday?

Hint. Ignore February 29 and assume that every person is equally likely to have been born on any of the 365 other days in a year.

Exercise 7 (Evaluation of the binomial distribution function in R) (5 points)
Let $X \sim \mathrm{B}(n, p)$, $n \in \mathbb{N}$, $p \in [0, 1]$. Show that the distribution function F of X satisfies

$$F(x) = \frac{1}{B(x+1, n-x)} \int_p^1 z^{(x+1)-1}(1-z)^{n-x-1}\,dz, \quad x \in \{0, \ldots, n-1\}, \tag{1}$$

where $B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1}\,dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ denotes the *Beta function.*
Note. This is how R evaluates `pbinom(x, size=n, prob=p)`.
Hint. View the left-hand side and the ride-hand side of (1) as functions in p and show equality by showing equality of the derivatives in p and, additionally, equality of the function values for one specific p.

Exercise 8 (M/M/1 queue events) (5 points)
Consider an M/M/1 queue with interarrival times 0.5, 0.6, 0.3, 0.9, 1.0, 0.8, 0.4 and service times 0.1, 0.2, 0.7, 0.4, 0.3, 0.5, 0.9. Give a snapshot of the system by completing the table below (using the notation and algorithm from class).

Simulation clock		System state			Event list		Statistical counters	
t	T	U_t	Q_t	$t_{a,q}$	t_a	t_d	n	TD_n
1	3.4	1	3	(0.2, 0.6, 1)	3.1	2	8	5.3

Exercise 9 (Implementation of an M/M/1 queue) (5 points)
Implement an M/M/1 queuing system in R with the following arguments:

- Input: Interarrival times `A`, service times `S`, number of delayed customers `n`
- Output: The delay (time between entering the queue and starting being served) of each customer

Run the code with:

```
set.seed(271) # set a seed
A <- rexp(1000, 6) # generate interarrival times
S <- rexp(1000, 2.5) # generate service times
n <- 400 # number of delayed customers
MM1(A, S=S, n=n) # call your implemented function MM1()
```

Compute the mean delay $\bar{D}_n$ based on all $n =$ `n` customers. As part of your solution, print your source code and include the R output for computing $\bar{D}_n$.