Assignment 3

1. A travelling salesman has carpets in his van. He goes from door to door trying to sell them. Each time he sells one he gets excited and is able to be more convincing and each time he is rejected and cannot close a deal he gets less excited and is less convincing when dealing with the next customer. This salesman was blessed with 10 levels of excitement and each interaction either lifts or decreases his excitement by one level. The probabilities for the salesman to sell a carpet depending on how he feels are {0, 0.1, 0.2, 0.3, …0.9}. Write an algorithm to simulate this Markov chain. In b) - f) run simulations for all initial levels of excitement, **except the first level**, because business is downright non-existent when one feels this low for starters. Use one bin per timestep. **Run simulations over the time period (sales events) of $T = 1000$. For statistics, run 10 000 simulations**, which in the present story means that the salesman makes 10 000 trips.
   a) (1 p) **Write down the Markov matrix.**
   b) (4 p) **For every simulation of each initial level of excitement, compute the total number of sold carpets by the end of the simulation (cumulative number of sales). For each initial level of excitement plot as a normalised histogram (density=True) the probability mass density (PMD) of the cumulative number of sales** (9 histograms in total)**.** To make this clear, such a histogram has the cumulative number of sales on the x-axis and the (normalised) frequency with which each number of cumulative sales is realised in 10000 trials on the y-axis. Choose the appropriate scale to see the PMD as clearly as possible.
   c) (2 p) For each different initial level, compute the average (over the 10 000 trials) number of carpets the salesman can sale within the given time T. **Plot these average numbers of carpets as a function of the salesman's initial level of excitement** (1 plot)**.**
   d) (4 p total) (3 p) **Compute and plot the PMFs for time (the number of steps) required for the salesman to sell 5 carpets. Exclude the runs that ended up without selling at least 5 carpets**. At sufficiently high initial levels of excitement the PMFs should start showing a definite functional form. (1 p) **Choose appropriate scale (axes) for plotting so that you can name this functional form – and name it. (Use this scale for all plots in d).)**
   e) (2 p) Compute the **average** times (for the different initial levels) it takes for the salesman to sell 5 carpets. Plot these times as a function of the salesman's initial level of excitement.
   f) (1 p) In point d) and e), excluding the runs that didn't sell enough carpets is underestimating the number of steps needed to sell 5 carpets. **Plot the probability of selling less than 5 carpets in T timesteps as a function of the initial level of excitement.**
   g) (1 p) In d) you obtained some definite functional form(s) for certain level(s) of initial excitement. **Why does the current process produce these functional forms? (**If you did not get these forms, you can still argue for them based on what you know.)

2. Let's take the Poisson process to describe radioactive decay. The number of nuclei is initially $N_0 = 10000$. The nuclei decay (fission) at rate $\lambda = 0.2$ per second. We want to

determine the half-time $t_{1/2}$, that is, the time it takes on average for the number of nuclei to decay to $N(t) = N_0/2$. You can do this simulation in two ways, the first of which is what a statistician would do and that is presented in Lecture 3. There is an alternative way based on simulating the stochastic process in time steps, which is what for example a physicist would do. In this second way you should first run simulations to find appropriate time interval (time step) but let's pretend you have already done this and found that $\Delta t = 0.01$ s.

a) (3 p) Implement the algorithm to simulate radioactive decay by both methods.
b) (2 p) Determine the mean and variance for $t_{1/2}$ that you get by averaging over 100 simulations.

Note! Use the definitions of the lecture notes. There exists an alternative definition for the parameters in the exponential distribution; do not use that. You should, of course get the same values for mean and variance by both methods. If you get a very large values (e.g. hundreds, thousands, …) for the mean and variance, you have probably used the alternative definition.

Hint - one way to think about this: The decay is exponential (Poisson process), so $N(t) = N_0 e^{-\lambda t}$. (So, you could calculate $t_{1/2}$ exactly, but the point here is to figure out how to simulate the process.) For each nucleus, **the probability not to decay in time *t*** is then $P = \frac{N}{N_0} = e^{-\lambda t}$. In the second "time-interval" way you can **think** in terms of **the probability of a decay per time step**. (See Lecture 3, page 35.)

Note 2! Although the probability of decay **per constant time interval and per nucleus** is constant, **the decay probability for an ensemble of nuclei** decreases with decreasing N. This is an **inhomogeneous** Poisson process: the differential equation describing the process is $\frac{dN(t)}{dt} = -\lambda N(t)$, so the transition probability changes with N and so with time t. Hence, the time-dependent rate of this process is $\lambda_{eff}(t) = \lambda N(t)$. Now inter-event time distribution is $Exp(\lambda_{eff}(t))$ and the number-of-event distribution is $N_t \sim Po(\lambda_{eff}(t))$. So, basically you need to make sure that $\lambda_{eff}(t)$ changes with time – and so N – as it should in your simulation. See also TA's slide in "assignment-3" in Slack.

Note 3 Any two methods, one using the given Δt and the other not using it will be regarded as correct in peergrade.

(Needless to say, write the algorithms for the Poisson process, that is, don't use some all-inclusive black-box library function – no points for that.)