J Kohonen

3B Distribution of sums, and normal approximation

Class problems

To calculate CDF values of a normal distribution, you can use a computer (see e.g. the R command `pnorm`), or use tables such as Mellin's tables provided on the course page, or Appendix A1 in Ross's book.

3B1 (Road salting) To keep the roads free from ice, the officials have stored enough salt for a total snowfall of 200 cm. (We assume, for simplicity, that the amount of salt needed is proportional to the snowfall.) The daily snowfalls have expectation 4.5 cm and standard deviation 2.5 cm.

(a) Using the normal approximation (central limit theorem), calculate an approximate probability that the salt stored will suffice for 50 days.

(b) What additional assumptions did you have to make to be able to solve (a)? Do you think the assumptions are realistic?

3B2 (Winning at the casino) A simplified roulette wheel has 37 slots, numbered from 0 to 36. On every round of the game, Harry bets one euro on his personal lucky number. If the ball lands on that number, he receives 36 euros (net gain +35 because he spent 1 eur on the bet). Otherwise, he loses the euro (net gain -1). All slots are equally likely. We denote Harry's net gain after n rounds by $S_n = X_1 + \cdots + X_n$, where X_i is the net gain from round i.

(a) Find the mean and standard deviation of X_i.

(b) Find the mean and standard deviation of S_n.

Applying the normal approximation (central limit theorem), calculate the approximate probability that Harry's net gain is positive

(c) after 30 rounds,

(d) after 3 000 rounds,

(e) after 300 000 rounds.

The first of these probabilities can also be calculated exactly, with relative ease:

(f) Calculate the probability in (c) exactly, and compare to the approximate value.
Hint: What exactly must happen so that Harry's net gain would be zero or negative, after 30 rounds?

Home problems

3B3 (Getting enough responses) Opinion pollsters have calculated (by some method) that they need responses from at least 100 people for their opinion poll, in order to have enough data for the statistical inference they wish to conduct. From previous experience, they estimate that when they send their poll to a person, the person has 80% probability of responding, independently from the other persons. Thus they decide to send the poll to 140 persons.

(a) Applying the binomial distribution (without normal approximation), find the probabilities that the pollsters get exactly x responses, for values $x = 100$, 112 and 120. Find the density function of binomial distribution from e.g. Ross's section 5.1. To calculate it, you can use a computer.

(b) Applying the normal approximation (central limit theorem), find approximately the probability that the pollsters get enough responses (at least 100). Use tables or a computer to calculate the CDF.

3B4 (Stock portfolios) The shares of two companies, Xanadu and Ypsilon, are on the stock market. Currently both are priced at 100 euros. The next-year return on one share of Xanadu (in euros) is modelled as a random variable X that follows normal distribution with mean $\mu_X = 15$ and standard deviation $\sigma_X = 10$. The return on one share of Ypsilon is modelled as Y, which follows normal distribution with parameters $\mu_Y = 10$ and $\sigma_Y = 10$. The returns of the two shares are assumed independent.

(a) Abel buys 200 shares of Xanadu, so he will gain $A = 200X$ during the next year. Find the distribution of A. What is the probability that Abel loses money (has negative gain)?

(b) Bertha buys 100 shares of Xanadu and 100 shares of Ypsilon, so she will gain $B = 100X + 100Y$. Find the distribution of B. What is the probability that Bertha loses money?

(c) Find the correlation of A and B. Are they independent?

(d) Find the distribution of $A - B$. What is the probability that Abel makes more money than Bertha?

Hint. The sum of two independent normally distributed variables is also normally distributed, but what are its parameters? Recall also how scaling (multiplication by constant) works. Note also that $-Y = (-1) \cdot Y$.