STA 247 — Answers for problem set #2

Question 1: The random variable X has the binomial distribution with parameters $n = 60$ and $p = 1/40$. The random variable Y has the binomial distribution with $n = 48$ and $p = 1/30$. Prove that $P(X + Y \geq 31)$ is no more than $1/10$.

We can prove this using Markov's inequality. X and Y are non-negative, so $X + Y$ is also non-negative. A binomial(n,p) random variable has expectation np, so $E(X) = 60/40 = 1.5$ and $E(Y) = 48/30 = 1.6$. Since $E(X + Y) = E(X) + E(Y)$, we see that $E(X + Y) = 3.1$. Markov's inequality then tells us that $P(X + Y \geq 31) \leq E(X + Y)/31 = 3.1/31 = 1/10$.

Question 2: You have been informed that the main U of T web page is accessed an average of 25000 times per day. You have also been told that this web page is accessed more than 50000 times on 1% of the days. Say whatever you can about the standard deviation of the number of accesses in a day.

Let X be the number of accesses in a day. We know that $\mu = E(X) = 25000$. We also know that $P(X > 50000) = 0.01$. Since the number of accesses in a day can't be negative, this is equivalent to $P(|X - 25000| > 25000) = 0.01$.

Using Chebychev's inequality, we see that

$$0.01 = P(|X-25000| > 25000) = P(|X-25000| \geq 25001) = P(|X-\mu| \geq 25001) \leq \sigma^2/25001^2$$

where σ is the standard deviation of X. From this we get that $\sigma^2 \geq 0.01 \times 25001^2$, and therefore $\sigma \geq 2500.1$.

We might be able to say something stronger about σ if we knew more about the distribution, but not if we have to rely only on Chebychev's inequality.

Question 3: Suppose we roll 10 fair six-sided dice. Let S be the sum of the numbers showing on all of these dice. Find the mean and standard deviation of S, and the mean and standard deviation of $S/10$, which is the average value shown on the 10 dice.

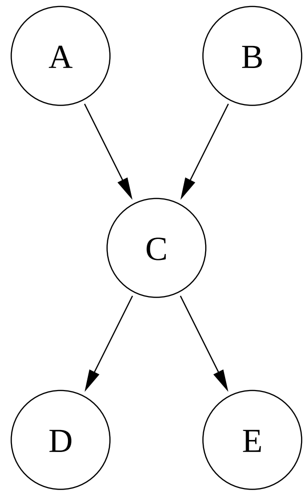
Let the numbers on the ten dice be $X_1, \ldots, X_{10}$.

Direct computaton gives $E(X_i) = 3.5$ and Var$(X_i) = 2.91666\ldots$ for all i.

From this, we get that $E(S) = E(X_1) + \cdots + E(X_{10}) = 10 \times 3.5 = 35$, and $E(S/10) = E(S)/10 = 3.5$.

Since the rolls of different dice will be independent, we have that Var(S) = Var$(X_1) + \cdots +$ Var$(X_{10}) = 10 \times 2.91666\ldots = 29.1666\ldots$, and hence SD$(S) = \sqrt{29.1666\ldots} = 5.400617\ldots$. We then get Var$(S/10)$ = Var$(S)/10^2 = 0.291666\ldots$ and SD$(S/10) = 0.5400617\ldots$.

Question 4: Suppose that the joint distribution of the random variables A, B, C, D, and E is described by the following directed graphical model:



Suppose also that the marginal distributions of A and B are both binomial(2,1/4), the conditional distribution of C given $A = a$ and $B = a$ is Bernoulli($(a + b)/4$), and the conditional distributions of D and E given $C = c$ are both Bernoulli($c/2$).

a) Compute $P(A = 1, B = 2, C = 1, D = 0, E = 1)$.

According to the directed graphical model,

$$\begin{aligned} &P(A = 1, B = 2, C = 1, D = 0, E = 1) \\ &\quad = P(A = 1)\, P(B = 2)\, P(C = 1|A = 1, B = 2)\, P(D = 0|C = 1)\, P(E = 1|C = 1) \\ &\quad = [2(1/4)(3/4)]\, [(1/4)(1/4)]\, [3/4]\, [1/2]\, [1/2] \\ &\quad = 0.00439453125 \end{aligned}$$

b) Find $P(A = 0, B = 0|C = 1)$.

$P(A = 0, B = 0|C = 1) = P(A = 0, B = 0, C = 1) \,/\, P(C = 1)$. But when $A = 0$ and $B = 0$, the condtional distribution of C given these values for A and B is Bernoulli(0), for which $C = 1$ has probability zero. Therefore $P(A = 0, B = 0|C = 1) = 0$.

One can calculate that $P(A = 0|C = 1)$ and $P(B = 0|C = 1)$ are both non-zero, and so their product is non-zero. This confirms that A is not conditionally independent of B given C, as one would suspect from the directed graphical model (though one cannot conclude this with certainty from the graph alone).

c) Find $P(D = 0, E = 0|C = 1)$.

According to the graphical model, D and E are conditionally independent given C, so

$$P(D = 0, E = 0|C = 1) = P(D = 0|C = 1)\, P(E = 0|C = 1) = (1/2)\,(1/2) = 1/4$$