Tentamentsskrivning i Sannolikhetsteori 1, MSG100, del 1, 7.5 hp.

Tid: Onsdagen den 20 oktober, 2010 kl 08.30-12.30
Examinator och jour: Serik Sagitov, tel. 772-5351, mob. 0736 907 613, rum H3026 i MV-huset.
Hjälpmedel: Miniräknare, **egen** formelsamling (4 sidor på 2 blad A4).
Betygränser: för "G" fordras 12 poäng, för "VG" - 20 poäng.
==

1. (5 points) Two events A and B have $P(A) = 0.3$, $P(B) = 0.7$ and $P(B|A) = 0.5$.

a. Draw a Venn diagram illustrating this situation.

b. Compute $P(A|B)$.

c. The corresponding indicators 1_A and 1_B are random variables with Bernoulli distributions. Without computing their correlation coefficient ρ make a guess about the sign of ρ. Explain your guess.

d. Compute the correlation coefficient ρ.

2. (5 points) Normal approximation for the Poisson distribution.

a. Clearly write down the statement on normal approximation for the Poisson distribution in terms of the cumulative distribution functions (fördelningsfunktioner). How does it follow from the Central Limit Theorem?

b. Give a direct proof of the fact that the Poisson probability mass function (sannolikhetsfunktion) is approximated by the normal density function.

3. (5 points) A single experiment involves one or two coin tossings. Toss a coin, if you see tails stop the experiment, otherwise toss another coin. Denote by X the number of heads. For 10 independent runs of the experiment denote by $X_1, \ldots, X_{10}$ the corresponding numbers of heads.

a. What is the exact meaning of the statement "random variables X_1 and X_2 are independent" in terms of their joint distribition and marginal distributions?

b. Draw a tree-diagram for the outcomes of the first two experiments. Indicate on the tips of the tree the outcomes constituting the event

$$B = \{X_1 + X_2 = 2\}.$$

What is $P(B)$?

c. Compute the probability of the event

$C = \{$four out of ten numbers $X_1, \ldots, X_{10}$ are equal to 0, and three of these numbers are equal to 2$\}$.

4. (5 points) The kurtosis of the distribution of X with mean μ and variance σ^2 is defined as

$$\gamma_4 = E\left(\frac{X-\mu}{\sigma}\right)^4.$$

a. Why is the transform $E(e^{tX})$ called the moment generating function of X?

b. Find the kurtosis of the normal distribution.

5. (5 points) Three random variables U_1, U_2, U_3 are independent and unoformly distributed $U(0,1)$ on the unit interval.

a. What is the joint density function $f(x_1, x_2, x_3)$ of the random vector (U_1, U_2, U_3)?

b. Let X be the maximum among U_1, U_2, U_3. Compute its cumulative distribution function.

c. Find the mean value (väntevärde) of X.

6. (5 points) Let X have an exponential distribution with parameter λ and put $Y = e^{-X}$.

a. Specify the interval of values that Y can take.

b. Compute the variance of Y.

c. Find the density of Y. It belongs to a well-known family of distributions - which one?

Partial answers and solutions are also welcome. Good luck!

COMPUTATIONAL ANSWERS

1b.
$$P(A|B) = \frac{0.5 \cdot 0.3}{0.7} = 0.21$$

1c. One should expect the negative sign of ρ since
$$P(A|B) < P(A), P(B|A) < P(B)$$
meaning negative dependence: given that one of the events occurs, the other event has smaller conditional probability compared to the unconditional probability.

1d. Since
$$E(1_A) = P(A) = 0.3, E(1_B) = P(B) = 0.7, E(1_A 1_B) = P(A \cap B) = 0.15,$$
we have
$$Cov(1_A, 1_B) = 0.15 - 0.3 \cdot 0.7 = -0.06.$$
Futhermore, since
$$Var(1_A) = Var(1_B) = 0.3 \cdot 0.7 = 0.21,$$
we obtain
$$\rho = \frac{-0.06}{0.21} = -0.29.$$

2b. Let $\lambda \to \infty$ and $k \to \infty$ in such a way that $x = \frac{k-\lambda}{\sqrt{\lambda}}$ is a constant. Then $k \sim \lambda$ and by the Stirling formula
$$\begin{aligned} P(X = k) &= \frac{\lambda^k}{k!} e^{-\lambda} \sim \frac{(\lambda/k)^k}{\sqrt{2\pi k}} e^{k-\lambda} \\ &\sim \frac{1}{\sqrt{2\pi\lambda}} \left(\frac{k - x\sqrt{\lambda}}{k} \right)^k e^{x\sqrt{\lambda}} \\ &= \frac{1}{\sqrt{2\pi\lambda}} e^{k \ln(1 - x\sqrt{\lambda}/k)} e^{x\sqrt{\lambda}}. \end{aligned}$$
Now since
$$k \ln(1 - x\sqrt{\lambda}/k) + x\sqrt{\lambda} \sim \frac{x^2 \lambda}{2k} \to x^2/2,$$
we conclude
$$P(X = k) \sim \frac{1}{\sqrt{2\pi\lambda}} e^{-x^2/2}.$$

3a. Since X_i take values 0,1,2, independence between X_1 and X_2 means
$$P(X_1 = i, X_2 = j) = P(X_1 = i) P(X_2 = j)$$
for $i, j = 0, 1, 2$.

3b.
$$P(B) = \frac{1}{2} \cdot \frac{1}{4} + \frac{1}{4} \cdot \frac{1}{4} + \frac{1}{4} \cdot \frac{1}{2} = \frac{5}{16}$$

3c. Using the multinomial distribution $\mathrm{Mn}(10, \frac{1}{2}, \frac{1}{4}, \frac{1}{4})$ we get
$$P(C) = \frac{10!}{4!3!3!} 2^{-4} 4^{-6} = 0.064$$

4a. From $M(t) = E(e^{tX})$ the k-th moment is computed by
$$E(X^K) = \frac{d^k}{dt^k} M(t)|_{t=0}$$

4b. If $X \sim N(\mu, \sigma^2)$, then $\gamma_4 = E(Z^4)$ for $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$. Since

$$E(e^{tZ}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-x^2/2} dx = e^{t^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x-t)^2/2} dx = e^{t^2/2}$$

we obtain

$$\begin{aligned} \gamma_4 &= \frac{d^4}{dt^4} e^{t^2/2}|_{t=0} = \frac{d^3}{dt^3}(te^{t^2/2})|_{t=0} = \frac{d^2}{dt^2}(e^{t^2/2} + t^2 e^{t^2/2})|_{t=0} \\ &= (3te^{t^2/2} + t^3 e^{t^2/2})'|_{t=0} = (3e^{t^2/2} + 6t^2 e^{t^2/2} + t^4 e^{t^2/2})|_{t=0} = 3 \end{aligned}$$

5a.

$$f(x_1, x_2, x_3) = \begin{cases} 1 & \text{if } 0 < x_1 < 1, 0 < x_2 < 1, 0 < x_3 < 1 \\ 0 & \text{otherwise} \end{cases}$$

5b.

$$P(X \le x) = P(U_1 \le x, U_2 \le x, U_3 \le x) = x^3, \ 0 \le x \le 1$$

5c.

$$E(X) = \int_0^1 x \cdot (x^3)' dx = 3 \int_0^1 x^3 dx = 3/4$$

6a.

$$Y \in (0, 1]$$

6b. After computing

$$E(Y) = \lambda \int_0^{\infty} e^{-x} e^{-\lambda x} dx = \frac{\lambda}{\lambda + 1}$$

and

$$E(Y^2) = \lambda \int_0^{\infty} e^{-2x} e^{-\lambda x} dx = \frac{\lambda}{\lambda + 2}$$

we conclude

$$Var(Y) = \frac{\lambda}{\lambda + 2} - \left(\frac{\lambda}{\lambda + 1} \right)^2 = \frac{\lambda}{(\lambda + 2)(\lambda + 1)^2}$$

6c. From

$$P(Y \le y) = P(X > -\ln y) = e^{\lambda \ln y} = y^{\lambda}$$

we derive the density of the Beta$(\lambda, 1)$ distribution

$$f(y) = (y^{\lambda})' = \lambda y^{\lambda - 1}.$$