Probability

Math 493 — Fall 2008

Suggested solutions — Final Exam, December 16th, 2008

Problem 2

The natural assumption is to assume that the lifetimes of the lightbulbs are exponentially distributed with parameter $\lambda = \frac{1}{170}$. Arguments could also be made for assuming that the lifetimes are either normally or uniformly distributed, although further assumptions need to be made on the variance. Note that it does *not* make sense to assume that the lifetimes are Poisson distributed.

Let $X_1, X_2, \ldots, X_{50}$ be the lifetimes of the 50 lightbulbs. Assuming that each X_i is exponentially distributed with parameter $\lambda = \frac{1}{170}$, we have

$$E[X_i] = \frac{1}{\lambda} = 170 \qquad \text{and} \qquad \operatorname{Var}(X_i) = \frac{1}{\lambda^2} = 28900.$$

By the Central Limit Theorem the sum $X_1 + X_2 + \cdots + X_{50}$ is approximately normally distributed with expected value $50 \cdot 170 = 8500$ and variance $50 \cdot 28900 = 1445000$. Therefore

$$\begin{aligned} P(X_1 + \cdots + X_{50} \geq 8760) &= P\left(\frac{X_1 + \cdots + X_{50} - 8500}{\sqrt{1445000}} \geq \frac{8760 - 8500}{\sqrt{1445000}} \right) \\ &\approx 1 - \Phi\left(\frac{8760 - 8500}{\sqrt{1445000}} \right) \approx 1 - \Phi\big(0.216\big) \\ &\approx 1 - 0.5855 = 0.4145. \end{aligned}$$

Problem 3

You should *not* switch. Switching gives a $\frac{1}{3}$ chance of winning a car, while not switching gives a $\frac{2}{3}$ chance.

To see this, define the events

G_i the goat is behind door i,

H_{ij} the host opens door j after the player has picked door i.

Then

$$P(G_1) = P(G_2) = P(G_3) = \tfrac{1}{3},$$

and

$$P(H_{ij} \mid G_k) = \begin{cases} 0 & \text{if } i = j \text{ (can not open the door the player picked)} \\ 0 & \text{if } j = k \text{ (can not open the door hiding the goat)} \\ \frac{1}{2} & \text{if } i = k \text{ (two doors equally likely to be opened)} \\ 1 & \text{if all } i, j, k \text{ are different (only one door possible to open).} \end{cases}$$

Assume (by relabeling the doors if necessary) that the player picks door 1 and that the host opens door 3. Then the goat is behind either door 1 or door 2. The player wins by not switching only if the goat is behind door 2, while he wins by switching only if it is behind door 1.

Thus, the probability of winning by *not* switching is

$$\begin{aligned} P(G_2 \mid H_{13}) &= \frac{P(H_{13} \mid G_2)P(G_2)}{P(H_{13} \mid G_1)P(G_1) + P(H_{13} \mid G_2)P(G_2) + P(H_{13} \mid G_3)P(G_3)} \\ &= \frac{1 \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3}} = \frac{2}{3}. \end{aligned}$$

The probability of winning by switching is

$$\begin{aligned} P(G_1 \mid H_{13}) &= \frac{P(H_{13} \mid G_1)P(G_1)}{P(H_{13} \mid G_1)P(G_1) + P(H_{13} \mid G_2)P(G_2) + P(H_{13} \mid G_3)P(G_3)} \\ &= \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3}} = \frac{1}{3}. \end{aligned}$$

Problem 4

Distribution	Density function	Exp. value	Variance
Uniform (a, b)	$f(x) = \frac{1}{b-a}, \quad a < x < b$	$\frac{1}{2}(a+b)$	$\frac{1}{12}(b-a)^2$
Normal (μ, σ)	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}, -\infty < x < \infty$	μ	σ^2
Exponential (λ)	$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

Problem 5

Write X as a sum of independent Bernoulli random variables, $X = X_1 + \cdots + X_n$, where

$$X_i = \begin{cases} 1 & \text{if the } i\text{th experiment is a success} \\ 0 & \text{otherwise.} \end{cases}$$

Then $E[X_i] = P(X_i = 1) = p$, and since $X_i^2 = X_i$, $E[X_i^2] = p$. Thus

$$\text{Var}(X_i) = E[X_i^2] - E[X_i]^2 = p - p^2.$$

It then follows that

$$E[X] = \sum_{i=1}^{n} E[X_i] = np,$$

and

$$\text{Var}(X) = \sum_{i=1}^{n} \text{Var}(X_i) = n(p - p^2) = np(1-p).$$

This can also be proved directly, as is done on pages 155 – 156 in the book.

Problem 6

Let X denote Michelle's score and let Y denote Barack's score. Then

$$P(Y > X) = P(Y - X > 0) = P(Y - X > 0.5) = P(Z > 0.5),$$

where $Z = Y - X$ and a continuity correction has been performed. The random variable Z is normally distributed with

$$E[Z] = E[Y - X] = E[Y + (-1)X] = E[Y] + (-1)E[X] = E[Y] - E[X] = 84 - 86 = -2,$$

and

$$\begin{aligned}\operatorname{Var}(Z) &= \operatorname{Var}(Y - X) = \operatorname{Var}(Y + (-1)X) = \operatorname{Var}(Y) + (-1)^2 \operatorname{Var}(X) \\ &= \operatorname{Var}(Y) + \operatorname{Var}(X) = 10^2 + 7^2 = 149.\end{aligned}$$

Thus,

$$\begin{aligned}P(Z > 0.5) &= P\left(\frac{Z - (-2)}{\sqrt{149}} > \frac{0.5 - (-2)}{\sqrt{149}}\right) = 1 - \Phi\left(\frac{2.5}{\sqrt{149}}\right) \\ &\approx 1 - \Phi(0.205) \approx 1 - 0.5812 = 0.4188.\end{aligned}$$

Problem 7

The important observation is the for each n, there are $n + 1$ ways the nonnegative integer valued variables X and Y sum up to n. Thus, the probability mass function of $X + Y$ is

$$\begin{aligned}P(X + Y = n) &= \sum_{k=0}^{n} P(X = k, Y = n - k) = \sum_{k=0}^{n} P(X = k)P(Y = n - k) \\ &= \sum_{k=0}^{n} e^{-\lambda_1} \frac{\lambda_1^k}{k!} \cdot e^{-\lambda_2} \frac{\lambda_2^{n-k}}{(n-k)!} = e^{-(\lambda_1+\lambda_2)} \sum_{k=0}^{n} \frac{1}{(n-k)!k!} \lambda_1^k \lambda_2^{n-k}.\end{aligned}$$

By the Binomial Theorem

$$\sum_{k=0}^{n} \binom{n}{k} \lambda_1^k \lambda_2^{n-k} = \sum_{k=0}^{n} \frac{n!}{(n-k)!k!} \lambda_1^k \lambda_2^{n-k} = \big(\lambda_1 + \lambda_2\big)^n.$$

Therefore,

$$P(X + Y = n) = e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1 + \lambda_2)^n}{n!},$$

and $X + Y$ is Poisson distributed with parameter $\lambda_1 + \lambda_2$.

Problem 8

The joint density of X and Y is

$$f(x,y) = \begin{cases} Cxy(x^2+y^2) & 0 \le x, y \le 10, \\ 0 & \text{otherwise.} \end{cases}$$

a) To find C, we use that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1.$$

As

$$\int_0^{10}\int_0^{10} xy(x^2+y^2)\,\mathrm{d}x\,\mathrm{d}y = \int_0^{10} 2500y + 50y^3\,\mathrm{d}y = 250000,$$

we have $C = \frac{1}{250000}$.

Next,

$$\begin{aligned} P(X > 2Y) &= \iint_{x>2y} f(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &= C\int_0^{10}\int_0^{x/2} xy(x^2+y^2)\,\mathrm{d}y\,\mathrm{d}x = C\int_0^{5}\int_{2y}^{10} xy(x^2+y^2)\,\mathrm{d}x\,\mathrm{d}y. \end{aligned}$$

Either integration gives $P(X > 2Y) = \frac{3}{32}$.

b) The conditional density function of X given $Y = y$ is

$$f_{X|Y}(x \mid y) = \frac{f(x,y)}{f_Y(y)} = \frac{Cxy(x^2+y^2)}{C\int_0^{10} xy(x^2+y^2)\,\mathrm{d}x} = \frac{Cxy(x^2+y^2)}{Cy(2500+50y^2)} = \frac{x(x^2+y^2)}{2500+50y^2}.$$

We can then calculate

$$\begin{aligned} P(4 < X < 9 \mid Y = 3) &= \int_4^9 f_{X|Y}(x \mid 3)\,\mathrm{d}x = \int_4^9 \frac{x(x^2+3^2)}{2500+50\cdot 3^2}\,\mathrm{d}x \\ &= \frac{1}{2950}\int_4^9 x^3 + 9x\,\mathrm{d}x = \frac{299}{472}. \end{aligned}$$

Problem 9

a) Let X be a random variable with finite expected value $E[X] = \mu$, and (finite) variance $\mathrm{Var}(X) = \sigma^2$. Then for any $k > 0$,

$$P\big(|X - \mu| \geq k\big) \leq \tfrac{\sigma^2}{k^2}.$$

To prove this, consider the nonnegative random variable $Y = (X - \mu)^2$, which has expected value $E[Y] = E[(X - \mu)^2] = \mathrm{Var}(X) = \sigma^2$. By Markov's inequality,

$$P\big(Y \geq k^2\big) = P\big((X - \mu)^2 \geq k^2\big) \leq \tfrac{\sigma^2}{k^2}.$$

Since $(X - \mu)^2 \geq k^2$ whenever $|X - \mu| \geq k$, also $P\big(|X - \mu| \geq k\big) \leq \tfrac{\sigma^2}{k^2}$.

b) Let $X_1, X_2 \ldots$ be a sequence of independent, identically distributed random variables, each having finite expected value $E[X_i] = \mu$ (and finite variance $\mathrm{Var}(X_i) = \sigma^2$). Then, for any $\epsilon > 0$,

$$\lim_{n \to \infty} P\big(\big|\tfrac{X_1 + \cdots + X_n}{n} - \mu\big| \geq \epsilon\big) = 0.$$

To prove this, note that for any n,

$$E\big[\tfrac{X_1 + \cdots + X_n}{n}\big] = \mu \qquad \text{and} \qquad \mathrm{Var}\big(\tfrac{X_1 + \cdots + X_n}{n}\big) = \tfrac{\sigma^2}{n}.$$

Then Chebyshev's inequality tells us that

$$P\big(\big|\tfrac{X_1 + \cdots + X_n}{n} - \mu\big| \geq \epsilon\big) \leq \tfrac{\sigma^2}{n\epsilon^2},$$

which tends to 0 as n tends to ∞.

Problem 10

Let X denote the number of occupied tables, and think of $X = X_1 + \cdots + X_N$, where

$$X_i = \begin{cases} 1 & \text{if the } i\text{th person sits at an unoccupied table} \\ 0 & \text{otherwise.} \end{cases}$$

Since the ith person sits at an unoccupied table only if he is not friends with any of the $i - 1$ previous people. Therefore,

$$E[X_i] = P(X_i = 1) = (1 - p)^{i-1}.$$

Summing up, the expected number of occupied tables is

$$E[X] = \sum_{i=1}^{N} E[X_i] = \sum_{i=1}^{N} (1 - p)^{i-1} = \tfrac{1-(1-p)^N}{1-(1-p)} = \tfrac{1-(1-p)^N}{p}.$$