Math/Stats 425, Sec. 1, Fall '04: Introduction to Probability

Final Exam: Solutions

1. In a game, a contestant is shown two identical envelopes containing money. The contestant does not know how much money is in either envelope. The contestant can choose an envelope at random (equally likely), see how much money is in it, and then either keep this amount or exchange the chosen envelope for the other one, keeping the money in the second envelope. Now suppose that one envelope has \$1000, and the other has \$2000. A strategy to play this game is for the contestant to fix a number x, and if the first envelope is revealed to have $\leq \$x$, then one switches; if it has more than $\$x$, one keeps the first envelope. What is the expected amount of money the contestant gets following this strategy if $x = 500$? What is the expected amount won if the strategy is used with $x = 1500$? Finally, what is the expected amount won if the strategy is used with $x = 2500$?

[25]

• This is a simple expectation calculation. Let $\mathbf{Y}$ be the RV "the winnings". The two possibilities are according to whether the \$1000 envelope or the \$2000 envelope is chosen, and these are equally likely. In the first strategy, $x = 500$, if one were to draw \$ 1000, one would hold it, and if one drew \$2000, one would hold that, too. So, the expectation under this strategy is $E(\mathbf{Y}|x = 500) = 1000 \cdot \frac{1}{2} + 2000 \cdot \frac{1}{2} = 1500$. When $x = 1500$, if one drew \$1000, one would exchange it for the \$2000 envelope, and if one drew \$2000, one keeps it. Therefore, $E(\mathbf{Y}|x = 1500) = 2000 \cdot \frac{1}{2} + 2000 \cdot \frac{1}{2} = 2000$. Finally, if one used the strategy for $x = 2500$, then if one drew \$1000, one would turn it back for the envelope with \$2000, and if one drew \$2000, one would exchange it for the envelope with \$1000, so $E(\mathbf{Y}|x = 2500) = 2000 \cdot \frac{1}{2} + 1000 \cdot \frac{1}{2} = 1500$ again. In fact,

$$E(\mathbf{Y}|x) = \begin{cases} 1500, & \text{if } x < 1000, \text{or } x \geq 2000, \text{and,} \\ 2000, & \text{if } 1000 \leq x < 2000. \end{cases}$$

Remark: Since you don't know beforehand what the amounts in the envelopes will be, one could randomize one's choice of the strategy x, i.e., replace x by a non-negative random variable $\mathbf{X}$ so that we could now compute the expected winnings with this strategy as

$$E(\mathbf{Y}) = E(E(\mathbf{Y}|\mathbf{X})),$$

by the conditional expectation formula, and this last is always ≥ 1500, and if $P(1000 \leq \mathbf{X} < 2000) > 0$, then $E(\mathbf{Y}) > 1500$. It pays to guess!

2. Let $\mathbf{Z}$ be the *unit normal* random variable, i.e., a normal R.V. with mean 0 and variance 1. Let $\mathbf{Y}$ be the random variable $\mathbf{Z}^2$. [$\mathbf{Y}$ is the χ^2 (*chi-squared*) distribution with one degree of freedom. The χ^2-distributions are very important in statistics.]

(i) What are the possible values of $\mathbf{Y}$?

• Since $\mathbf{Z}$ can take on any real value, $\mathbf{Y}$ can take on any non-negative real value.

(ii) What is the formula for the probability density function $f_{\mathbf{Y}}(y)$?

• We use the usual technique: $f_{\mathbf{Y}}(y) = \frac{d}{dy} F_{\mathbf{Y}}(y)$, and

$$F_{\mathbf{Y}}(y) = P(\mathbf{Y} \le y) = P(-\sqrt{y} \le \mathbf{Z} \le \sqrt{y}) = \frac{1}{\sqrt{2\pi}} \int_{-\sqrt{y}}^{\sqrt{y}} e^{-\frac{t^2}{2}} dt,$$

for $y \ge 0$. Taking the derivative of this integral with respect to y we get

$$\frac{d}{dy} F_{\mathbf{Y}}(y) = \frac{1}{\sqrt{2\pi}} \{ \frac{1}{2\sqrt{y}} e^{-y/2} - \frac{-1}{2\sqrt{y}} e^{-y/2} \},$$

where the first term inside the {}'s comes from differentiating the upper end point of the integral and the second term comes from differentiating the lower endpoint of the integral, using the fundamental theorem of calculus and the chain rule. Altogether we have

$$f_{\mathbf{Y}}(y) = \frac{1}{\sqrt{2\pi y}} e^{-y/2}, \text{for } 0 < y.$$

(iii) $\mathbf{Y}$ has the same distribution as one of the *Gamma* distributions. For which parameters (s, λ) is the Gamma distribution the same as that of $\mathbf{Y}$?

• The Gamma$[\lambda, s]$ distrbution has pdf, for $0 \le y$, given by $\frac{\lambda e^{-\lambda s} (\lambda y)^{s-1}}{\Gamma(s)}$. Comparing the exponents on the e's above, we see that $-\frac{y}{2}$ should match $-\lambda y$, which gives $\lambda = \frac{1}{2}$. Now comparing the exponents on the y's in both expressions we get $y^{-1/2} = y^{s-1}$, which gives $s = \frac{1}{2}$. All that is left to match is the constant $\Gamma(s)$, and the law of total probability then tells us that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Note that this last part is not a part of the question: that is, once you have matched the s and λ, and have only constant factors left to compare, the law of total probability tells you that they *have to match*, and it isn't necessary to show this to know that the probability density functions are the same.

[30]

3. Suppose $\mathbf{X}, \mathbf{Y}$ are two jointly distributed RV's with joint probability density function

$$f_{\mathbf{X},\mathbf{Y}}(x, y) = \begin{cases} 12xy(1-x) & 0 < x < 1, 0 < y < 1, \\ 0 & \text{otherwise} \end{cases}$$

(i) Check that $f_{\mathbf{X},\mathbf{Y}}(x, y)$ is indeed a probability density function.

• This just means checking two things: first, that $f_{\mathbf{X},\mathbf{Y}}(x, y)$ is non-negative for (x, y) in the square $S = \{0 < x < 1, 0 < y < 1\}$, which is obvious, and that $\int_S f_{\mathbf{X},\mathbf{Y}}(x, y)\, dx\, dy = 1$, which is true and is an elementary calculus exercise to check.

(ii) What is the probability density function of $\mathbf{X}$? What is the p.d.f. of $\mathbf{Y}$?

• To find the marginals, we just have to integrate out the y or the x according to the case. So,

$$f_{\mathbf{X}}(x) = \int_0^1 12xy(1-x)dy = 12x(1-x) \int_0^1 ydy = 6x(1-x), 0 < x < 1.$$

and

$$f_{\mathbf{Y}}(y) = 12y \int_0^1 x(1-x)dx = 2y, 0 < y < 1.$$

(iii) What is the conditional p.d.f. $f_{\mathbf{X}|\mathbf{Y}=y}(x)$?

• You can do this two ways. The first is to notice that since the p.d.f. above for $(\mathbf{X}, \mathbf{Y})$ factors as a function of x times a function of y, the two RV's are independent, and hence

$$f_{\mathbf{X}|\mathbf{Y}=y}(x) = f_{\mathbf{X}}(x).$$

Of course, you can also just make the calculation directly, too:

$$f_{\mathbf{X}|\mathbf{Y}=y}(x) = \frac{f_{\mathbf{X},\mathbf{Y}}(x,y)}{f_{\mathbf{Y}}(y)} = \frac{12xy(1-x)}{2y} = 6x(1-x) = f_{\mathbf{X}}(x).$$

(iv) Find the probability $P(\mathbf{Y} < \frac{1}{2}|\mathbf{X} > \frac{1}{2})$.

• Again, if you notice independence, we get:

$$P(\mathbf{Y} < \frac{1}{2}|\mathbf{X} > \frac{1}{2}) = P(\mathbf{Y} < \frac{1}{2}) = \int_0^{\frac{1}{2}} 2y\, dy = \frac{1}{4}.$$

or you can do it out again:

$$P(\mathbf{Y} < \frac{1}{2}|\mathbf{X} > \frac{1}{2}) = \frac{P(\mathbf{Y} < \frac{1}{2} \text{ and } \mathbf{X} > \frac{1}{2})}{P(\mathbf{X} > \frac{1}{2})} = \frac{\int_0^{\frac{1}{2}} \int_{\frac{1}{2}}^1 12xy(1-x)\, dx\, dy}{\int_{\frac{1}{2}}^1 6x(1-x)\, dx} = \frac{\frac{1}{8}}{\frac{1}{2}} = \frac{1}{4}.$$

[35]

4. Suppose we have two urns labeled A and B. A contains 3 maize marbles and 5 blue marbles; B contains seven maize marbles and 4 blue. A biased coin is flipped, for which $P(H) = \frac{2}{3}$. If the coin comes up heads, we draw a marble from urn A, and if it comes up tails, we draw a marble from urn B. What is the probability of drawing a blue marble? If you are told that the marble drawn was a yellow one, what is the probability that the coin came up heads when flipped?

• The first part is computing a probability by conditioning on cases, the cases being whether the head H was flipped or the tail T. Let b be the event "a blue marble was drawn", and m the event "a maize marble was drawn". You are first asked to find $P(b)$. This is simply:

$$P(b) = P(b|H)P(H) + P(b|T)P(T) = \frac{5}{8} \cdot \frac{2}{3} + \frac{4}{11} \cdot \frac{1}{3} = \frac{71}{131}.$$

Next you are asked to find the conditional probability of the coin having come up heads, if a maize marble was drawn. This is an example of Bayes's theorem:

$$P(H|m) = \frac{P(m|H)P(H)}{P(m|H)P(H) + P(m|T)P(T)} = \frac{\frac{3}{8} \cdot \frac{2}{3}}{\frac{3}{8} \cdot \frac{2}{3} + \frac{7}{11} \cdot \frac{1}{3}} = \frac{33}{61}.$$

[25]

5. (i) In general, if two random variables $\mathbf{U}_1, \mathbf{U}_2$ have the same moment generating functions, i.e.,

$$M_{\mathbf{U}_1}(t) = E(e^{t\mathbf{U}_1}) = E(e^{t\mathbf{U}_2}) = M_{\mathbf{U}_2}(t),$$

then what can you say about the RV's $\mathbf{U}_1, \mathbf{U}_2$?

• In this case, under technical assumptions which we will ignore here, the two distributions have to be the same: see the textbook, section 7.1, p. 366, just above example 6e, and examples 6e through 6h.

Now let $\mathbf{X}, \mathbf{Y}$ be two independent RV's both $\sim$ Exponential[λ]. Recall that these may be thought of as waiting time distributions.

(ii) Use moment generating functions to show that $\mathbf{X}+\mathbf{Y} \sim$ Gamma$[2, \lambda]$. What is the interpretation of this in terms of waiting times?

• Use the principle in part (i):

$$M_{\mathbf{X}+\mathbf{Y}}(t) = M_{\mathbf{X}}(t) \cdot M_{\mathbf{Y}}(t) = \frac{\lambda}{\lambda - t} \cdot \frac{\lambda}{\lambda - t} = (\frac{\lambda}{\lambda - t})^2,$$

where we have used the tables for moment generating functions. Again using the table for the m.g.f. for the Gamma$[s, \lambda]$ distribution, we see that the m.g.f. is $(\frac{\lambda}{\lambda - t})^s$, and so we have matched a Gamma distribution with $s = 2$ and λ as given.

In terms of waiting tiimes, it says that we can interpret Gamma[2,λ] as the waiting time distribution, waiting for the arrival of first one, then another event according to independent Exp[λ] distributions.

(iii) If $\mathbf{X}+\mathbf{Y} = T > 0$, what are the possible values of $\mathbf{X}$? What is the conditional distribution for $\mathbf{X}$, given that $\mathbf{X}+\mathbf{Y} = T$? Interpret this in terms of waiting times.

• In this conditioned situation, $0 \leq \mathbf{X} \leq T$, since $\mathbf{Y} \geq 0$. Then the conditional distribution function is given, for $0 \leq x \leq T$, by:

$$f_{\mathbf{X}|\mathbf{X}+\mathbf{Y}=T}(x) = \frac{f_{\mathbf{X},\mathbf{Y}}(x, T-x)}{f_{\mathbf{X}+\mathbf{Y}}(T)} = \frac{\lambda^2 e^{-\lambda x} \cdot e^{-\lambda(T-x)}}{\lambda e^{-\lambda T}(\lambda T)^{(2-1)}},$$

where the last numerator comes from the Gamma[2, λ] distribution. Doing a little algebra to cancel stuff top and bottom simplifies things to:

$$f_{\mathbf{X}|\mathbf{X}+\mathbf{Y}=T}(x) = \frac{1}{T}, \text{for } 0 \leq x \leq T,$$

i.e., $\mathbf{X}$ is unformly distributed over the interval $[0, T]$. This means , in terms of waiting times, that if we know the second of these two arrivals occurs at time T, then the first arrival was a random time anywhere prior to time T, i.e., uniformly distributed on the interval $[0, T]$.

Now, I will accept an argument like the above, but strictly speaking, you should make a change of variables first. I will show you the argument in detail, and point out why you get the right answer either way in this case (something = 1, so you can't see the possible difference).

Let us change (random) variables from $\mathbf{X}, \mathbf{Y}$, to $\mathbf{X}, \mathbf{S} = \mathbf{X}+\mathbf{Y}$. At the level of possible values, this amounts to $(x, s) = g(x, y) = (x, x+y) = (g_1(x, y), g_2(x, y))$. The change of variables formula gives us

$$f_{\mathbf{X},\mathbf{S}}(x, s) = f_{\mathbf{X},\mathbf{Y}}(g^{-1}(x, s)) \cdot |\det(Dg(g^{-1}(x, s))|,$$

where

$$Dg(x,y) = \begin{pmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial y} \\ \\ \frac{\partial g_2}{\partial x} & \frac{\partial g_2}{\partial y} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \\ 1 & 1 \end{pmatrix},$$

and so, $\det(Dg(g^{-1}(x,s)) = 1$, and

$$f_{\mathbf{X},\mathbf{S}}(x,s) = f_{\mathbf{X},\mathbf{Y}}(x, s-x) = \lambda e^{-\lambda x} \lambda e^{-\lambda(s-x)} = \lambda^2 e^{-\lambda s}, 0 \le x \le s.$$

We also have

$$f_{\mathbf{S}}(s) = \lambda e^{-\lambda s} (\lambda s)^{2-1} = \lambda^2 s e^{-\lambda s}.$$

Now,

$$f_{\mathbf{X}|\mathbf{X}+\mathbf{Y}=T}(x) = f_{\mathbf{X}|\mathbf{S}=T}(x) = \frac{f_{\mathbf{X},\mathbf{S}}(x,T)}{f_S(T)} = \frac{\lambda^2 e^{-\lambda T}}{\lambda^2 T e^{-\lambda T}} = \frac{1}{T}, 0 \le x \le T,$$

as found above.

[30]

6. A certain component is necessary for the operation of a computer server system and must be replaced immediately upon failure for the server to continue operating. The lifetime of such a component is a random variable $\mathbf{X}$ with mean lifetime 100 hours and standard deviation 30 hours. Estimate the number of such components which must be in stock in order to be 95% sure that the system can run continuously for the next 2000 hours? (We ignore any other possible failure for the server!)

• Let us denote $\mathbf{X}_i$ a sequence of i.i.d.'s, all with the same distribution as $\mathbf{X}$, and let $\mathbf{S}_N = \mathbf{X}_1 + \ldots + \mathbf{X}_N$ be the total time one can run the server if one has N copies of the critical component. The question asks you to estimate an N for which $P(\mathbf{S}_N \ge 2000) \ge 0.95$. We can use the Central Limit Theorem to do this. Let's check the expectation of $\mathbf{S}_N$:

$$E(\mathbf{S}_N) = NE(\mathbf{X}_1) = 100N.$$

To estimate this, we are thinking of N so large that 2000 is an outlier on the short side, so we are probably looking for an N significantly > 20. In terms of the CLT, we want:

$$0.95 \le P(\mathbf{S}_N \ge 2000) = P(\frac{\mathbf{S}_N - 100N}{30\sqrt{N}} \ge \frac{2000-100N}{30\sqrt{N}}) \approx P(\mathbf{Z} \le \frac{100N - 2000}{30\sqrt{N}}),$$

and since the cumulative distribution function $F_{\mathbf{Z}}(z) = \Phi(z)$, we have to choose N at least large enough to make $\frac{100N-2000}{30\sqrt{N}} \ge 1.645$, since $\Phi(1.645) = 0.95$. That is, we want to guarantee that

$$100N - 1.645 \cdot 30\sqrt{N} - 2000 \ge 0.$$

Now replacing $\sqrt{N}$ by t (and N by t^2), we want to guarantee that

$$100t^2 - 1.645 \cdot 30t - 2000 \ge 0.$$

Now you can either do the next step by trial and error or a calculator, or just say what should happen at this point. I would try to solve these equations for the first value of $t > 0$ where the parabola crosses the x-axis, that is, solve

$$100t^2 - 1.645 \cdot 30t - 2000 = 0.$$

The positive root is $t = 4.726$, or $N \ge 22.332$. So, according to this estimate, we have to have at least 23 components to get the components working at least 2000 hours, with probability $\ge 95\%$.

[25]

7. Ten champions are competing in a team hot dog eating contest, representing the US, Japan, Mexico and Italy. Of the ten, 4 are from the US, 2 are from Japan, 1 is from Mexico and 3 are from Italy. The outcome of the contest is compiled only according to the national team, but not noting the individual competitors' names. How many possible results are there for this competition? How many possible results are there which have two Italians in the top three, and the third Italian among the last three?

• This is a combinatorics problem. The answer for how many lists there are is just a multinomial coefficient. More explicitly, the results are lists of contestants ordered according to how many hot dogs they ate. This gives us 10! possible lists, if we remember the names of the contestants. We are told to forget that, so we have to divide 10! for overcounting. Thus, we have to divide by 4! for the US contestants, and 2! for Japanese contestants, etc. The final answer is

$$N = \frac{10!}{4! \cdot 2! \cdot 1! \cdot 3!} = \binom{10}{4, 2, 1, 3}.$$

Now if we also have to account for the conditions described in the second part involving the Italian team, we distinguish them at first. We have to choose two of the first three slots for them (which is $\binom{3}{2} = 3$ possibilities and independently choose $\binom{3}{1} = 3$ possible slots for the Italian in the last three in the list. For each such way of laying out the three Italian positions, there are 7 remaining slots. So we can make $3 \cdot 3 \cdot 7!$ lists which consist of individual USers, Japanese and Mexican, and three Italian slots. We have to correct for the fact that the US, Japan and Mexico are still individual names, meaning we have to divide as before by $4! \cdot 2! \cdot 1!$, i.e.,

$$N = 3 \cdot 3 \cdot \binom{7}{4, 2, 1} = \frac{3 \cdot 3 \cdot 7!}{4! \cdot 2! \cdot 1!} = 945.$$

[30]