NAME: (print!) Dr. Z.

E-Mail address: DrZprob at gmail dot com

Solutions to MATH 477 (3), Dr. Z. , Exam 2, Monday, Nov. 27, 2017, 8:40-10:00am, HLL 116

EXPLAIN EVERYTHING! Only simple calculators are allowed. You need the Z table.

1. (10 pts.) Let X, Y, Z be three random variables for which

$$Var(X) = 1 \quad , \quad Var(Y) = 1 \quad , \quad Var(Z) = 1 \quad ,$$

$$Var(X+Y) = 4 \quad , \quad Var(X+Z) = 4 \quad , \quad Var(Y+Z) = 4 \quad .$$

Find $Var(X+Y+Z)$

---

ans. 9

Before we can do anything else, we need $Cov(X,Y)$, $Cov(X,Z)$, and $Cov(Y,Z)$. By the famous formula

$$Var(X+Y) = Var(X) + Var(Y) + 2\,Cov(X,Y) \quad ,$$

we have

$$4 = 1 + 1 + 2\,Cov(X,Y) \quad .$$

Hence, $Cov(X,Y) = (4-1-1)/2 = 1$.

Similarly, $Cov(X,Z) = 1$ and $Cov(Y,Z) = 1$.

Now we use, the three-r.v. version of the above famous formula

$$Var(X+Y+Z) = Var(X) + Var(Y) + Var(Z) + 2\,Cov(X,Y) + 2\,Cov(X,Z) + 2\,Cov(Y,Z) \quad ,$$

getting

$$Var(X+Y+Z) = 1 + 1 + 1 + 2\cdot 1 + 2\cdot 1 + 2\cdot 1 = 1+1+1+2+2+2 = 9 \quad .$$

2. (10 points altogether)

(i) (3 points) State the formula for the expected number of rounds it takes for a gambler who enters a fair casino with x dollars, until she is is either broke or has N dollars. At each round she wins a dollar with probability $\frac{1}{2}$ and loses a dollar with probability $\frac{1}{2}$.

---

ans. $x(N-x)$

(ii) (7 points) Give a complete proof, with all the details, of that formula. (You may use , w/o proof, the fact that the system of equations has a unique solution)

Let $f(x)$ be the expected duration remaining at the casino if you currently have x dollars. Suppose that $1 \leq x \leq N-1$.

There are two possibilities, each with probability $\frac{1}{2}$, as follows.

• You lost a dollar, so now you have $x-1$ dollars, and the expected **remaining** duration is, by definition, $f(x-1)$. So the expected remaining duration, **conditioned** on losing is $\frac{1}{2} \cdot f(x-1)$.

• You won a dollar, so now you have $x+1$ dollars, and the expected **remaining** duration is, by definition, $f(x+1)$. So the expected remaining duration, **conditioned** on winning is $\frac{1}{2} \cdot f(x+1)$.

Adding these two contributions, the expected remaining duration, after the next round is $\frac{1}{2}f(x-1) + \frac{1}{2}f(x+1)$, and adding the current round, 1, we have the $N-1$ equations

$$f(x) \, = \, \frac{1}{2}f(x-1) \, + \, \frac{1}{2}f(x+1) \, + \, 1 \quad , \quad (1 \leq x \leq N-1) \quad .$$

In addition, we have the obvious two equations

$$f(0) = 0 \quad , \quad f(N) = 0 \quad ,$$

since if you are broke you are immediately out, and if you have N dollars, you are also immediately out.

Let

$$g(x) = x(N-x) \quad .$$

Then we claim that

$$g(x) \, = \, \frac{1}{2}g(x-1) \, + \, \frac{1}{2}g(x+1) \, + \, 1 \quad .$$

In other words

$$g(x) - \frac{1}{2}g(x-1) - \, \frac{1}{2}g(x+1) - 1 \, = \, 0 \quad .$$

Simple algebra yields, that indeed

$$x(N-x) - \frac{1}{2}(x-1)(N-x+1) - \frac{1}{2}(x+1)(N-x-1) - 1 \;=\; 0 \quad .$$

Also

$$g(0) = 0 \cdot (N-0) = 0 \quad , \quad g(N) = N \cdot (N-N) = 0 \quad .$$

Hence **both** $(f(0), f(1), \cdots, f(N))$ and $(g(0), g(1), \cdots, g(N))$ are solutions to the **same** system of $N+1$ equations and $N+1$ unknowns. By **uniqueness**, they must be the same. Hence $f(x) = x(N-x)$. **QED**.

3. (10 points altogether)

(i) (3 points) State the formula for the expected number of coupons that a coupon-collector must buy before he gets a full collection, and there are N coupons altogether. It is assumed that at each purchase each of these coupons is equally likely to show up (but of course, you only see what you got after you open the package).

---

ans. $N\left(1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{N}\right)$ or $N\sum_{i=1}^{N}\frac{1}{i}$.

(ii) (7 points) Using the fact that the Expectation of a Geometric random variable with parameter p is $\frac{1}{p}$, and the Geometric random variable is the number of trials until the first success, when you independently try something whose probability is p, prove the above formula.

Let T be the r.v. "total number of coupons bought until one has **all** N kinds". We can break-up T into a sum of N simpler random variables, for $0 \leq i \leq N-1$.

T_i: The number of coupons bought after the collector has i **different** kinds, until he gets yet-another-kind (i.e. until he has $i+1$ **different** kinds).

Then of course

$$T = \sum_{i=0}^{N-1} T_i \quad .$$

By **Linearity of Expectation**

$$E[T] = \sum_{i=0}^{N-1} E[T_i] \quad .$$

But each T_i is a **Geometric** r.v. with probability of success $\frac{N-i}{N}$, since if you currently have i different kinds, there are $N-i$ kinds that you **don't** have, hence the probability that if you buy a coupon, it would be brand-new is $\frac{N-i}{N}$. Since, as is well-known (and you are allowed to use), the expectation of a geometric r.v. with parameter p is $\frac{1}{p}$, we have

$$E[T_i] = \frac{N}{N-i} \quad .$$

Hence

$$E[T] = \sum_{i=0}^{N-1} \frac{N}{N-i} == N\sum_{i=0}^{N-1}\frac{1}{N-i} == N\sum_{i=1}^{N}\frac{1}{i} \quad .$$

QED.

4. (10 points) If you enter a casino with 800 dollars, and wish to make 1000 dollars, and the probability, at each round, of winning a dollar is 0.49 and losing a dollar is 0.51, what is the probability of exiting a loser?

---

ans. $0.9996648951\ldots$

We use the formula that the probability of **winning** is

$$\frac{1-(q/p)^x}{1-(q/p)^N} \quad .$$

Here $p = 0.49$, $q = 1 - p = 0.51$, $x = 800$ and $N = 1000$, so the probability of winning is

$$\frac{1-(0.51/0.49)^{800}}{1-(0.51/0.49)^{1000}} \quad ,$$

and hence the probability of **losing** is

$$1 - \frac{1-(0.51/0.49)^{800}}{1-(0.51/0.49)^{1000}} = 0.9996648951\ldots \quad .$$

5. (10 points) The joint density function of X and Y is given by

$$f(x,y) = \begin{cases} \frac{x+y}{5} \quad , & if \quad 0 < x < 2\,,\, 1 < y < 2\,; \\ 0 \quad , & otherwise, \end{cases}$$

If you know that $Y = 1.25$ what is the probability that $1 \leq X \leq 2$.

---

ans. $\frac{11}{18} = 0.6111111111\ldots$

The desired probability is the **quotient**

$$\frac{\int_1^2 f(x, \frac{5}{4})\,dx}{\int_0^2 f(x, \frac{5}{4})\,dx} \quad .$$

Hence, in this problem

$$\frac{\int_1^2 \frac{x+\frac{5}{4}}{5}\,dx}{\int_0^2 \frac{x+\frac{5}{4}}{5}\,dx} \quad .$$

Multiplying top and bottom by 20 (to make the computation more user-friendly), this is

$$\frac{\int_1^2 (4x+5)\,dx}{\int_0^2 (4x+5)\,dx} \quad .$$

The bottom is

$$\int_0^2 (4x+5)\,dx \;=\; (2x^2+5x)\Big|_0^2 = \; 2(2^2-0^2)+5(2-0) \;=\; 8+10=18 \quad .$$

The top is

$$\int_1^2 (4x+5)\,dx \;=\; (2x^2+5x)\Big|_1^2 = \; 2(2^2-1^2)+5(2-1) \;=\; 6+5=11 \quad .$$

Hence the desired probability is $\frac{11}{18} = 0.6111111111111111\ldots$.

6. (10 points) In a certain community of married couples, the maximal income of the wife is 100K and the maximal income of the husband is also 100K. Every husband makes **at most** what his wife makes. Let X denote the the wife's income and let Y denote the husband's income. Let X and Y have joint density function $f(x,y) = 3\,(x^2+y^2)$ on the region where the density is positive. The unit of money is 100K.
If it is known that the wife makes 50K dollars, what is the probability that the husband makes less that 25K?

---

ans. $\frac{13}{32} = 0.40625$

The desired probability is the **quotient**

$$\frac{\int_0^{\frac{1}{4}} f(\frac{1}{2}, y)\, dy}{\int_0^{\frac{1}{2}} f(\frac{1}{2}, y)\, dy}$$

Hence, in this problem

$$\frac{\int_0^{\frac{1}{4}} 3\,((\frac{1}{2})^2 + y^2)\, dy}{\int_0^{\frac{1}{2}} 3\,((\frac{1}{2})^2 + y^2)\, dy}$$

Canceling out the 3, and multiplying top and bottom by 4 (to make the computation more user-friendly), this is

$$\frac{\int_0^{\frac{1}{4}} (1+4y^2)\, dy}{\int_0^{\frac{1}{2}} (1+4y^2)\, dy}$$

The bottom is

$$\int_0^{\frac{1}{2}} (1+4y^2)\, dy \;=\; \left(y + 4\frac{y^3}{3}\right)\Big|_0^{\frac{1}{2}} = (\frac{1}{2} - 0) + \frac{4}{3}((\frac{1}{2})^3 - 0^3) = \frac{1}{2} + \frac{1}{6} = \frac{2}{3} \quad .$$

The top is

$$\int_0^{\frac{1}{4}} (1+4y^2)\, dy \;=\; \left(y + 4\frac{y^3}{3}\right)\Big|_0^{\frac{1}{4}}$$
$$= (\frac{1}{4} - 0) + \frac{4}{3}((\frac{1}{4})^3 - 0^3) = \frac{1}{4} + \frac{1}{3}\cdot\frac{1}{4^2} = \frac{1}{4}\cdot(1 + \frac{1}{12}) = \frac{13}{48} \quad .$$

Hence the desired probability is $\frac{\frac{13}{48}}{\frac{2}{3}} = \frac{13}{32} = 0.40625$.

7. (10 points) Let X and Y be continuous random variables with joint density function

$$f(x,y) = \begin{cases} 30x & , \quad for \quad x^4 \le y \le x^3 \,; \\ 0 & , \quad otherwise \end{cases} \quad .$$

Let g be the marginal density function of X. Find $g(x)$.

---

ans.

$$g(x) \;=\; \begin{cases} 30x^4 - 30x^5 & , \quad if \quad 0 \le x \le 1 \,; \\ 0 & , \quad otherwise \end{cases} \quad .$$

Since the curves $y = x^3$ and $y = x^4$ intersect when $x^3 - x^4 = 0$, i.e. $x^3(x-1) = 0$, i.e. $x = 0$ and $x = 1$, the projection of the region on the x-axis is the interval $0 \le x \le 1$.

Hence, the region where the joint pdf is non-zero is

$$\{(x,y) \,|\, 0 \le x \le 1 \quad , \quad x^4 \le y \le x \,\} \quad .$$

The desired margical distribution, $f_X(x)$, is, when $0 \le x \le 1$,

$$\int_{x^4}^{x^3} 30x \, dy \;=\; 30x \int_{x^4}^{x^3} 1 \, dy \;=\; 30x \, (y) \Big|_{x^4}^{x^3} = 30x(x^3 - x^4) = 30x^4 - 30x^5 \quad .$$

Of course, it is 0 otherwise.

8. (10 points) A piece of equipment is being insured against early failure. The time from the purchase until failure of the equipment is exponentially distributed with mean 5 years. The insurance will pay an amount of $2A$ if the equipment fails during the first year, and it pays A if failure occurs during the second, third, or fourth year. If failure occurs after the first four years, no payment will be made.
At what level must A be set if the expected payment made under this insurance is to be 2000?

---

ans. $\frac{2000}{2-e^{-1/5}-e^{-4/5}} = 2732.463354\ldots$

The **parameter** is $\lambda = \frac{1}{5}$, hence the pdf is $f(x) = \frac{1}{5}e^{-x/5}$. The expected payment is

$$2A \int_0^1 \frac{1}{5}e^{-x/5}\,dx \,+\, A\int_1^4 \frac{1}{5}e^{-x/5}\,dx \,=$$

$$2A\left(-e^{-x/5}\right)\Big|_0^1 + A\left(-e^{-x/5}\right)\Big|_1^4 \,=\, 2A(-e^{-1/5} - -e^0) \,+\, A(-e^{-4/5} - -e^{-1/5})$$

$$= \, A(2 - 2e^{-1/5} + e^{-1/5} - e^{-4/5}) \,=\, A(2 - e^{-1/5} - e^{-4/5}) \quad .$$

Setting this equal to 2000, we get the equation

$$A(2 - e^{-1/5} - e^{-4/5}) \,=\, 2000 \quad .$$

Hence

$$A = \frac{2000}{2 - e^{-1/5} - e^{-4/5}} = 2732.463354\ldots \quad .$$

9. (10 points) Approximate the probability that if you toss a loaded coin, with $Pr(Head) = 0.6$, one thousand times , the number of times it lends Heads is ≥ 560 and ≤ 620.

---

ans. $\Phi(\frac{20.5}{\sqrt{240}}) - \Phi(\frac{-40.5}{\sqrt{240}}) = \Phi(1.323269310) - \Phi(-2.614263759) = 0.9026560843\ldots$.

We need

$$P\{560 \leq X \leq 620\} \quad ,$$

where X is a **binomial** r.v. with **parameters** $n = 1000$ and $p = 0.6$. As is well-known, $\mu = n \cdot p = 1000 \cdot 0.6 = 600$, and $\sigma^2 = np(1-p) = 1000 \cdot 0.6 \cdot 0.4 = 240$.
We **approximate** it by a **Normal** r.v. with the above μ and σ^2, **but**, we use the **half-integer** convention

$$P\{559.5 \leq X \leq 620.5\} \quad ,$$

Subtracting μ from everywhere, we need

$$P\{559.5 - 600 \leq X - 600 \leq 620.5 - 600\} \quad ,$$

i.e.

$$P\{-40.5 \leq X - 600 \leq 20.5\}$$

Dividing by $\sigma = \sqrt{240}$ everywhere, we need

$$P\{\frac{-40.5}{\sqrt{240}} \leq \frac{X-600}{\sqrt{240}} \leq \frac{20.5}{\sqrt{240}}\}$$

So

$$P\{\frac{-40.5}{\sqrt{240}} \leq Z \leq \frac{20.5}{\sqrt{240}}\} \quad .$$

This is

$$\Phi(\frac{20.5}{\sqrt{240}}) - \Phi(\frac{-40.5}{\sqrt{240}}) = \Phi(1.323269310) - \Phi(-2.614263759) = 0.9026560843\ldots \quad .$$

10. (10 points)
Out of a class of 80 students,
• 20 students play football and soccer
• 20 students play football and basketball
• 20 students play soccer and basketball
• 10 students play football soccer and basketball
• 10 students play none of these three sports
• The number of students who play each of the three sports is the same
How many students play soccer?

---

ans. 40

Let x be the number of students who play soccer. By the last clue, the number of students who play basketball is also x, and the number of students who play football is also x. Hence

$$|B| = x \quad , \quad |F| = x \quad , \quad |S| = x .$$

We also know from the data that

$$|BF| = 20 \quad , \quad |BS| = 20 \quad , \quad |SF| = 20 \quad ,$$

and

$$|BFS| = 10 \quad , |B^c F^c S^c| = 10 \quad , |U| = 80 \quad ,$$

where U is the **universal set**, of all students.

By the famous **Principle of Inclusion-Exclusion**

$$|B^c F^c S^c| \;=\; |U| - (|B| + |F| + |S|) + (|BF| + |BS| + |FS|) - |BFS| \quad .$$

Hence

$$10 = 80 - (x + x + x) + (20 + 20 + 20) - 10 \quad ,$$

hence

$$10 = 80 - 3x + 60 - 10 \quad ,$$

hence

$$3x = 120 \quad ,$$

hence $x = 40$. Hence the number of students who play soccer is 40.