Lecture Stat 302
Introduction to Probability - Slides 23

AD

April 2010

Exercise 1

- Alf and Beth are two UBC students. They must take 4 300-level courses from a list of 12 possibilities. If they select their courses independently and at random, what is the probability that they will have exactly two courses in common?
- Each student has $\binom{12}{4}$ possible choices. So the total number of choices for both students is $A = \binom{12}{4}^2 = 495^2$. The number of choices such that exactly 2 courses are in common is given by

$$B = \binom{12}{2} \times \binom{10}{2} \times \binom{8}{2} = 83160$$

so the probability is

$$P = \frac{B}{A} = 0.3394.$$

Exercise 2

- Suppose that 10% of the children are left-handed.
- (a) In a class of 20 children, what is the probability that at least two are left-handed?
- (b) Suppose a school has 10 classes of 20 children. If you check the classes one by one, what is the probability that the first left-handed child will be found in the fourth class?

Exercise 2

- (a) Let X denote the number of left-handed children in a class of 20 children. X follow a binomial distribution of parameter $n = 20$, $p = 0.1$ so

$$\begin{aligned} P(X \geq 2) &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - (1-p)^{20} - \binom{20}{1}(1-p)^{19} p \\ &= 0.6083 \end{aligned}$$

- (b) Let α be the probability that there is at least one left-handed child in a class of 20, then

$$\alpha = P(X > 0) = 1 - (1-p)^{20} = 0.8784.$$

  Let Y denote the number of classes you need to check to find a left-handed child, Y follows a geometric distribution of parameter α so

$$P(Y = 4) = (1-\alpha)^3 \alpha = 0.0016.$$

Exercise 3

- To estimate the number of trout in a lake, we caught 50 trout, tagged them and released them back in the lake. Later, we caught 40 trout and found out that 4 of them were tagged. From this experiment, estimate N, the number of trout in the lake. (Hint: Let p_N be the probability that, in a lake with N trout, exactly 4 of the 50 trout caught are tagged. Find the value of N that maximizes p_N).
- Let X the number of trout recaptured, then X follows an hypergeometric distribution of parameters N, $m = 50$ and $n = 40$. We have

$$P(X = 4) = p_N = \frac{\binom{50}{4}\binom{N-50}{36}}{\binom{N}{40}}$$

- This proba is maximized with respect to N at the point $\widehat{N} \approx mn/4 = 500$. (see book and lecture)

Exercise 4

- In answering a multiple choice exam question, a student either knows the correct answer or randomly pick 1 of m alternatives. Let p be the probability that the student knows the answer. If the student gets the correct answer, what is the probability that the student actually knew the answer?
- Let $A =$"know correct answer" and $B =$"give correct answer". We are interested in

$$P\left(A|B\right) = \frac{P\left(B|A\right)P\left(A\right)}{P\left(B\right)}$$

  where

$$P\left(B\right) = P\left(B|A^c\right)P\left(A^c\right) + P\left(B|A\right)P\left(A\right).$$

- We have $P\left(A\right) = 1 - P\left(A^c\right) = p$, $P\left(B|A\right) = 1$ and $P\left(B|A^c\right) = \frac{1}{m}$ so

$$P\left(A|B\right) = \frac{p}{p + \left(1-p\right)/m}.$$

Exercise 5

- Let X denote the random lifetime of a TV (in years) and assume that the density takes the form

$$f_X(x) = \begin{cases} \frac{1}{2}\exp(-x/2) & \text{for } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- (a) Calculate the proba that exactly 3 out of 4 TVs last more than 4 years.
- (b) Use both the change of variables formula AND the method of the distribution function to find the density of $Y = \sqrt{X}$.

Exercise 5

- (a) We have for a random lifetime X

$$\begin{aligned} P(X>4) &= \int_4^\infty f_X(x)\,dx = \frac{1}{2}\int_4^\infty \exp(-x/2)\,dx \\ &= \exp(-2) = 0.1353 \end{aligned}$$

- We want the proba that exactly 3 out of 4 last more than 4 years which is

$$\begin{aligned} P &= \left(\begin{array}{c} 4 \\ 3 \end{array}\right) [P(X>4)]^3 [P(X<4)] \\ &= 0.0086. \end{aligned}$$

Exercise 5

- (b) You can use the generic formula for $Y = g(X)$ where g is monotonic

$$f_Y(y) = f_X\left(g^{-1}(y)\right)\left|\frac{dx}{dy}\right|_{x=g^{-1}(y)} = f_X\left(g^{-1}(y)\right)\frac{1}{g'\left(g^{-1}(y)\right)}$$

where $y = g(x) = \sqrt{x}$ and $x = g^{-1}(y) = y^2$ so $g'\left(g^{-1}(y)\right) = \frac{1}{2y}$ and

$$f_Y(y) = 2yf_X\left(y^2\right)$$

- If we use the cdf, we have for $y \geq 0$

$$\begin{aligned} P(Y \leq y) &= P\left(\sqrt{X} \leq y\right) = P\left(X \leq y^2\right) \\ &= F_X\left(y^2\right) \end{aligned}$$

so by the chain's rule

$$f_Y(y) = 2yf_X\left(y^2\right)$$

Exercise 6

- Consider an exponential r.v. X of parameter λ; i.e. $f(x) = \lambda \exp(-\lambda x)$ for $x \geq 0$.
- (a) Compute $P(X \leq x | X \geq t)$ for $x, t \geq 0$.
- (b) Differentiate $P(X \leq x | X \geq t)$ w.r.t. to x f to obtain the conditional pdf $f_{X|X \geq t}(x)$.

Exercise 6

- (a) We have $P(X \leq x | X \geq t) = 0$ if $x \leq t$ and for $x \geq t \geq 0$

$$P(X \leq x | X \geq t) = \frac{P(X \leq x \cap X \geq t)}{P(X \geq t)} = \frac{P(t \leq X \leq x)}{P(X \geq t)}$$

where

$$\begin{aligned} P(t \leq X \leq x) &= \int_t^x \lambda \exp(-\lambda u)\, du = \exp(-\lambda t) - \exp(-\lambda x), \\ P(X \geq t) &= \int_t^\infty \lambda \exp(-\lambda u)\, du = \exp(-\lambda t). \end{aligned}$$

It follows that

$$P(X \leq x | X \geq t) = \frac{\exp(-\lambda t) - \exp(-\lambda x)}{\exp(-\lambda t)} = 1 - \exp(-\lambda(x - t))$$

- (b) By differentiation

$$f_{X|X \geq t}(x) = \lambda \exp(-\lambda(x - t)).$$

Hence $X - t$ is an exponential density of parameter λ, this is called the memoryless property of exponential.

Exercise 7

- The duration of a certain computer game is a random variable with an exponential distribution with mean 10 minutes. Suppose that when you enter a video arcade two players have just started playing the game (independently).
- (a) What is the probability that at least one of them is still playing 20 minutes later?
- (b) If we denote T_i the duration of the game of player i ($i = 1, 2$), what is the pdf of $T = \max(T_1, T_2)$? (Hint: compute first the cdf of T).

Exercise 7

- (a) T_1 and T_2 are independent exponential r.v. of parameter $\lambda = 1/20$ and we are interested in computing

$$\begin{aligned} & P(\{T_1 > 20\} \cup \{T_2 > 20\}) \\ = \; & P(T_1 > 20) + P(T_2 > 20) - P(\{T_1 > 20\} \cap \{T_2 > 20\}) \\ = \; & P(T_1 > 20) + P(T_2 > 20) - P(T_1 > 20)\, P(T_2 > 20) \end{aligned}$$

- We have

$$\begin{aligned} P(T_1 > 20) &= P(T_2 > 20) = \frac{1}{10}\int_{20}^{\infty} \exp(-t/10)\, dt \\ &= \exp(-2) = 0.1353 \end{aligned}$$

so

$$P(\{T_1 > 20\} \cup \{T_2 > 20\}) = 0.2524$$

Exercise 7

- (b) As suggested by the hint, we first compute the cdf of $T = \max(T_1, T_2)$

$$\begin{aligned}\Pr(T \leq t) &= \Pr(\max(T_1, T_2) \leq t) = \Pr(T_1 \leq t \cap T_2 \leq t) \\ &= \Pr(T_1 \leq t) \Pr(T_2 \leq t) \text{ (independence)}\end{aligned}$$

where

$$\Pr(T_1 \leq t) = \frac{1}{10} \int_0^t \exp(-u/10)\, du = 1 - \exp(-t/10)$$

- Hence we obtain

$$\Pr(T \leq t) = [1 - \exp(-t/10)]^2$$

and the pdf is obtained by differentiating the cdf

$$f_T(t) = \frac{1}{5} \exp(-t/10)(1 - \exp(-t/10)).$$

Exercise 8

- Two contractors, A and B, bid independently on a job. The contract will go to the lowest bidder. A's bid is a random number X selected with an uniform distribution on the interval $[0, 1]$, while B's bid Y has probability density $f_Y(y) = 2y$ for $0 < y < 1$ and $f_Y(y) = 0$ otherwise.
- (a) What is the joint pdf of X and Y?
- (b) What is the probability that A will win the contract, i.e. that $X < Y$?
- (c) What is the expected value of the winning bid $E[\min(X, Y)]$? (Hint: $\min(X, Y) = X$ for $X < Y$ and Y if $Y < X$).

Exercise 8

- (a) We have

$$f(x,y) = f_X(x) f_Y(y) = \begin{cases} 1 \cdot 2y = 2y & \text{for } 0 < x < 1,\ 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

- (b) We have

$$\begin{aligned} \Pr(X < Y) &= \int\int_{\{x,y:x<y\}} f(x,y)\,dxdy \\ &= \int_0^1 \left(\int_0^y f(x,y)\,dx \right) dy = 2\int_0^1 y^2 dy = \frac{2}{3} \end{aligned}$$

Exercise 8

- (c) We have

$$
\begin{aligned}
E\left[\min(X,Y)\right] &= \int\int \min(x,y)\cdot f(x,y)\,dxdy \\
&= \int\int_{\{x,y:x<y\}} x\cdot f(x,y)\,dxdy + \int\int_{\{x,y:x>y\}} y\cdot f(x,y)\,dxdy \\
&= \int_0^1 \left(\int_0^y x\cdot f(x,y)\,dx\right) dy + \int_0^1 \left(\int_0^x y\cdot f(x,y)\,dy\right) dx
\end{aligned}
$$

- Now

$$\int_0^1 \left(\int_0^y x\cdot f(x,y)\,dx\right) dy = \int_0^1 y^3 dy = \frac{1}{4}$$

and

$$\int_0^1 \left(\int_0^x y\cdot f(x,y)\,dy\right) dx = \frac{2}{3}\int_0^1 x^3 dx = \frac{2}{12} = \frac{1}{6}$$

so

$$E\left[\min(X,Y)\right] = \frac{1}{4} + \frac{1}{6}.$$