Alex Barnett
1/27/06

Math 50: Midterm 1 — SOLUTIONS

65 minutes, 70 points, no algebra-capable calculators. Show working/reaonsing, since only that way could you get partial credit. In the multiple-part questions, later parts are usually independent of earlier ones, so skip over one you can't do and come back later.

1. [16 points] An urn contains 4 red and 6 white chips. 5 chips are drawn at random, without replacement.

2 (a) What is the probability that the first three chips follow exactly the sequence RWR?

$$p(R)\,p(W|R)\,p(R|RW) = \frac{4}{10}\cdot\frac{6}{9}\cdot\frac{3}{8} = \frac{1}{10}$$

3 (b) What is the probability that the second chip is red?

marginalize (sum) over possible 1st chips.

$$p(2^{nd}R) = p(2^{nd}R|1^{st}W)\,p(1^{st}W) + p(2^{nd}R|1^{st}R)\,p(1^{st}R)$$
$$= \frac{4}{9}\cdot\frac{6}{10} + \frac{3}{9}\cdot\frac{4}{10} = \frac{4}{10}$$

← note that it's same as $p(1^{st}R)$!

3 (c) What is the probability that the first chip was red given that the second chip is red?

conditional: $A = 1^{st}R$, $B = 2^{nd}R$ (events)

$$p(A|B)\,p(B) = p(A\cap B) \quad \text{so } p(A|B) = \frac{p(A\cap B)}{p(B)} = \frac{\frac{4}{10}\cdot\frac{3}{9}}{4/10} = \frac{1}{3}$$

Note: time ordering irrelevant! (so you can use $p(B|A)$... easier!).

3 (d) What is probability of drawing a total of 2 red chips out of the 5?

Hypergeometric

$$p(2R \text{ in total}) = \frac{\binom{4}{2}\binom{6}{3}}{\binom{10}{5}} = \frac{\frac{4!}{2!2!}\,\frac{6!}{3!3!}}{\frac{10!}{5!5!}}$$

$$= \frac{6\cdot 20}{4\cdot 9\cdot 7} = \frac{10}{21} = 0.476\ldots$$

3 (e) Answer the previous question if they are drawn *with replacement.*

Now each draw is independent 'coin toss' with $p = \frac{4}{10}$ (chance of R)

⇒ Binomial with $n=5$, $k=2$ $\quad p_X(k) = \binom{5}{k} p^k (1-p)^{5-k}$

$$p_X(2) = \binom{5}{2}\left(\frac{2}{5}\right)^2\left(\frac{3}{5}\right)^3 = 10\cdot\frac{4\cdot 27}{5^5} = 0.3456$$

2 (f) What is the variance of the number of red chips drawn if they are drawn *with replacement*?

$$Var(X) = np(1-p) \quad \text{for binomial}$$
$$= 5 \cdot \frac{2}{5} \cdot \frac{3}{5} = \frac{6}{5}$$

2. [8 points] Your burglar alarm is 99% reliable (if someone is breaking into your house, this is the chance of it going off). However there is a 1% chance of it going off on a given night when there's no break-in. Police estimate that break-ins occur at a given house about 1 in 1000 nights. If you hear the alarm, what's the chances there's a break-in?

Define events A = alarm goes off
B = break-in

told
$$p(A|B) = 0.99$$
$$p(A|\bar{B}) = 0.01$$
$$p(B) = 0.001$$

so $p(A) = p(A|B)p(B) + p(A|\bar{B})p(\bar{B})$
$$= 0.99 \cdot 10^{-3} + 0.01 \cdot 0.999$$
$$= 0.01098$$

Bayes rule
$$p(B|A) = \frac{p(A|B)}{p(A)} p(B)$$
$$= \frac{0.99}{0.01098} 10^{-3} \approx 0.09 \approx \frac{1}{11}$$

not very likely it's a real break-in.

3. [16 points] Random variables X and Y are sampled from the joint pdf $f_{X,Y}(x,y) = c(2x+y)$, for $0 \leq X \leq 1$ and $0 \leq Y \leq 1$, for some constant c.

3 (a) Find c.

$$1 = \int_0^1\int_0^1 c(2x+y)\,dx\,dy = c\underbrace{\int_0^1\int_0^1 2x\,dx\,dy}_{x^2|_0^1 = 1} + c\underbrace{\int_0^1\int_0^1 y\,dx\,dy}_{\frac{1}{2}y^2|_0^1 = \frac{1}{2}} = \frac{3}{2}c$$

since integral must be 1 for valid pdf.

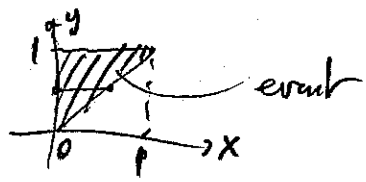
So $\boxed{c = \frac{2}{3}}$

4 (b) Find the marginal pdfs $f_X(x)$ and $f_Y(y)$.

$$f_X(x) = \int_0^1 f_{X,Y}(x,y)\,dy = \frac{2}{3}\int_0^1 (2x+y)\,dy = \frac{2}{3}\left[2xy + \frac{y^2}{2}\right]_{y=0}^1 = \frac{4x}{3} + \frac{1}{3}$$

$$f_Y(y) = \int_0^1 f_{X,Y}(x,y)\,dx = \frac{2}{3}\int_0^1 (2x+y)\,dx = \frac{2}{3}y + \frac{2}{3}x^2\Big|_0^1 = \frac{2y}{3} + \frac{2}{3}$$

double integral over a triangle.



4 (c) What is the probability that Y exceeds X?

$$P(Y > X) = \frac{2}{3}\int_0^1 \underbrace{\int_0^y (2x + y)\,dx}_{x^2 + xy\big|_{x=0}^{y} = y^2 + y^2 = 2y^2}\,dy$$

$$= \frac{2}{3}\int_0^1 2y^2\,dy = \frac{2}{3}\cdot\frac{2}{3}y^3\Big|_0^1 = \frac{4}{9}.$$

not $\frac{1}{2}$ since they're correlated (not indep.).

2 (d) Are X and Y independent? (Explain)

No, since $\underbrace{f_{X,Y}(x,y)}_{\text{joint}} \neq \underbrace{f_X(x)\,f_Y(y)}_{\text{product of marginals.}}$

3 (e) Find the expected value of Y given that X takes the value 1.

conditional $f_{Y|X}(y) = \dfrac{f_{X,Y}(x,y)}{f_X(x)} = \dfrac{\frac{2}{3}(2x+y)}{\frac{1}{3}(4x+1)} = 4\dfrac{x}{4x+1} + \dfrac{2}{4x+1}y$

$= \frac{4}{5} + \frac{2}{5}y$ when $x=1$.

$$E(Y|x) = \int_0^1 y\, f_{Y|x}(y)\,dy = \frac{1}{5}\int_0^1 y(4+2y)\,dy$$

$$= \frac{1}{5}\left[2y^2 + \frac{2}{3}y^3\right]_0^1 = \frac{8}{15}$$

4. [15 points] Each day you go to the Novack Cafe and buy (and eat) a bag of chips. According to the manufacturer, the weight (in ounces) of chips X in any bag is a random variable with pdf $f_X(x) = e^{-x}$, $x > 0$. You may leave your answers as formulae involving e if you wish. [Hint: $\int_0^\infty x^n e^{-x}dx = n!$].

3 (a) Find the pdf of the total weight of chips you ate in 2 days.

$W = X_1 + X_2$ Sum of 2 samples. (indep ones).

$$f_W(w) = \int_{-\infty}^{\infty} f_X(x)\, f_X(w-x)\,dx$$

for x and $w-x$ to be ≥ 0 need limits $0 \leq x \leq w$.

$$= \int_0^w e^{-x} e^{-(w-x)}\,dx = e^{-w}\int_0^w \cancel{e^{-x}}\,\cancel{e^{x}}\,dx = we^{-w}$$

for $w \geq 0$

(vanishes for $w < 0$)

3 (b) What is the expected total weight of chips eaten in 1 week (7 days)?

Let $W = X_1 + X_2 \cdots + X_7$, so $E(W) = E(X_1) + E(X_2) + \cdots E(X_7) = 7E(X)$

$E(X) = \int_0^{\infty} x e^{-x} dx = 1! = 1$ ← using formula I gave   so $E(W) = 7(1) = 7$

3 (c) What is the standard deviation of the total weight eaten in 1 week?

$Var(W) = Var(X_1) + \cdots + Var(X_7)$ since indep.

$= 7 \cdot Var(X)$ ← $Var(X) = E(X^2) - \mu^2 = \int_0^{\infty} x^2 e^{-x} dx - 1 = 1.$ $2!$ by formula.

$\sigma = \sqrt{Var(W)} = \sqrt{7}$.

3 (d) What are the chances that the smallest bag that week will *exceed* 1 ounce?

$p(\text{smallest of 7 samples} > 1) = p(\text{each sample} > 1) = p(X_1 > 1)p(X_2 > 1) \cdots p(X_7 > 1)$

$= (1 - F_X(1))^7 = (e^{-1})^7 = e^{-7}$ note $F_X(x) = \int_0^x e^{-y} dy = 1 - e^{-x}$

3 (e) What is the pdf of the weight of the largest bag that week?

Use formula. $f_{X'_{max}}(x) = n F_X(x)^{n-1} f_X(x) = 7(1-e^{-x})^6 e^{-x}$

(I apologize I gave the formula mistakenly as $F_{Y'_i} = \cdots$ not $f_{Y'_i} = \cdots$
(Note $F_{Y'_{max}}$ is actually a lot easier than this!) Therefore I removed no points for associated errors.

5. [15 points] In a small village there is a 1% probability of a birth occurring each day (assume this is independent from day to day, and constant).

3 (a) What is the mean number of births per year? (365 days)

mean each day $= 0.01$; means add, so

$\mu = 365 \times 0.01 = 3.65$ births/yr.

5 (b) Use a Poisson distribution to approximate the probability that 2 or more babies are born in a given half-year period.

rate $= \nu = 3.65$ Time interval $T = \frac{1}{2}$ (years)

so $\lambda = \nu T$

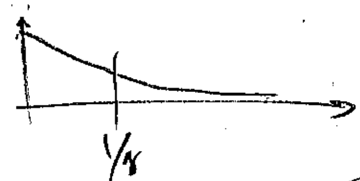
$p(2 \text{ or more}) = 1 - p(0) - p(1)$.

$= 1 - e^{-\lambda}\frac{\lambda^0}{0!} - e^{-\lambda}\frac{\lambda^1}{1!} = 1 - e^{-\lambda}(1+\lambda)$

≈ 0.5446.

3 (c) Let Y be the time (measured continuously in units of years) to the next birth. What is $f_Y(y)$?

rate γ



pdf of waiting time is $f_Y(y) = \gamma e^{-\gamma y}$

$= 3.65\, e^{-3.65 y}$

This is standard result for poisson process, continuous in time (we may approximate births as such since they are v. unlikely on a given day, and we consider time periods much longer than 1 day).

4 (d) The year after a chemical factory moves to town, no births are reported in an entire year. How concerned are you? Explain your reasoning. [Hint: how likely is this to happen presuming no change in underlying birth rate?]

$T = 1$ yr.

P(no births in 1 yr assuming fixed rate $\gamma = 3.65$)

$$= e^{-\gamma T} \frac{(\gamma T)^0}{0!} = e^{-3.65} \approx 0.026 \approx \frac{1}{38}$$

(it would happen every 38 yrs anyway.)

Therefore it seems a little unlikely this occurred by chance, but this does not provide compelling evidence that the birth rate γ has been decreased by the influence of chemical emissions.

One of you pointed out it was beyond the 0.05 'p value' used for hypothesis testing, though.