Practice problems # 3

Sta 230: Probability

April 4, 2017

This is a closed-book exam so do not refer to your notes, the text, or any other books (please put them on the floor). You may use the extra sheets discussed in in class.

You must show your work to get partial credit. Even correct answers will not receive full credit without justification.

Write your solutions as clearly as possible and make sure it's easy to find your answers (circle them if necessary), since you will not receive credit for work that I cannot understand or find. Good Luck!

If you find a question confusing please **ask me** to clarify it.

Problem 1 : [40 points] The joint pdf for the lifetime of two computer parts is

$$f(x, y) = xe^{-x(1+y)}, \quad x \geq 0, y \geq 0.$$

a. What is there marginal distribution for $f(x)$ and $(f(y)$?

b. Are the lifetimes of the two components independent ?

c. What is $f(x \mid y)$?

d. What is $\mathbf{P}(x + y < 2)$?

e. What is $\mathbf{P}(x > 4 \mid y = 3)$?

e. What is the probability that at least one of the components is lives longer than three years?

Problem 2 : [50 points] The result of an experiment testing a drug on mice was that 33 of the 67 mice recovered from the disease with a variance of .07%.

A clinical trial was then run on humans with this drug and 475 out of 502 patients were cured.

[1.] State the mean and variance of success for the mice ?

[2.] Given the mean and variance computed above use them to compute the parameters α and β of a Beta distribution.

[3.] Using the Beta distribution above as a prior compute the conditional density of success in the human trial

$$f(p \mid k = 475, n = 502)$$

[4.] State $\mathbf{E}[p \mid k = 475, n = 502]$ and $\mathbf{E}[p]$ before running the human trial.

[5.] Was the mouse experiment a useful study ?

Problem 3 : [40 points] Consider the following model

$$Y_i = 4 + \beta X_i + \varepsilon_i, \quad \varepsilon_i \overset{iid}{\sim} \mathrm{N}(0,1),$$

Determine

1. The conditional density for $Y \mid X$?
2. The pdf $f(y_1, ..., y_n \mid x_1, ..., x_n, \beta,)$.
3. Assume β is unknown, maximize the above pdf to provide an estimate for β.

Now consider the following model

$$Y_i = \exp(\beta X_i - 5 + \varepsilon_i), \quad \varepsilon_i \overset{iid}{\sim} \mathrm{N}(0,1),$$

Determine

3. The pdf $f(y_1, ..., y_n \mid x_1, ..., x_n, \beta,)$.

Problem 4: . [40 points] A factory has two types of machines. For each piece of furniture is made with either one of the machines. Machine one drills holes with the distribution $X_i \overset{iid}{\sim} \mathrm{N}(10, .1)$ and machine two with $X_i \overset{iid}{\sim} \mathrm{N}(12, .4)$.

1. You get a piece of furniture and measure 30 holes of size $(X_1, ..., X_{30})$ in your furniture. State the furniture the part you got was made from machine one and the probability the part you got was from machine two ? A piece of furniture goes to the either machine with equal probability.
2. State the four conditional probabilities, two marginal probabilities, and the one joint probability in this problem.
3. What is $\mathbf{E}[X]$.

(Nearly) Blank Work-Sheet #1 (of 8)

(Nearly) Blank Work-Sheet #2 (of 8)

(Nearly) Blank Work-Sheet #3 (of 8)

(Nearly) Blank Work-Sheet #4 (of 8)

(Nearly) Blank Work-Sheet #5 (of 8)

(Nearly) Blank Work-Sheet #6 (of 8)

(Nearly) Blank Work-Sheet #7 (of 8)

(Nearly) Blank Work-Sheet #8 (of 8)

Distributions

Binomial	$\Pr[X=x]$	$=$	$\binom{n}{x}p^x q^{(n-x)},\ x=0,1,\dots,n$	$\mu=np,\ \sigma^2=npq$
Poisson	$\Pr[X=k]$	$=$	$\lambda^k e^{-\lambda}/k!,\ k=0,1,\dots$	$\mu=\lambda,\ \sigma^2=\lambda$
Poisson process	$\Pr[X=k\vert t]$	$=$	$(\alpha t)^k e^{-\alpha t}/k!,\ k=0,1,\dots$	
Geometric	$\Pr[X=x]$	$=$	$pq^x,\ y=0,1,2,\dots$	$\mu=q/p,\ \sigma^2=q/p^2$
Hypergeometric	$\Pr[X=x]$	$=$	$\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}},\ x=0,1,2,\dots$	$\mu=n\frac{M}{N}$
				$\sigma^2=\frac{N-n}{N-1}n\frac{M}{N}(1-\frac{M}{N})$
Exponential	$\Pr[X\le x]$	$=$	$1-e^{-\lambda x},\ x\in(0,\infty)$	$\mu=1/\lambda,\ \sigma^2=1/\lambda^2$
	$f(x)$	$=$	$\lambda e^{-\lambda x},\ x>0;\ 0,\ x\le 0$	
Uniform	$\Pr[X\le x]$	$=$	$(x-a)/(b-a),\ x\in(a,b)$	$\mu=\frac{a+b}{2},\ \sigma^2=\frac{(b-a)^2}{12}$
	$f(x)$	$=$	$\frac{1}{b-a},\ a<x<b;\ 0,\ x\notin(a,b)$	
Normal	$\Pr[X\le x]$	$=$	$\Phi(\frac{x-\mu}{\sigma}),\ x\in\mathbb{R}$	$\mu=\mu,\ \sigma^2=\sigma^2$
	$f(x)$	$=$	$\frac{1}{\sqrt{2\pi}\sigma}e^{-(x-\mu)^2/2\sigma^2},\ x\in\mathbb{R}$	

$$P_k^n = \frac{n!}{(n-k)!} = \overbrace{(n)(n-1)\cdots(n-k+1)}^{k\text{ terms}}$$

$$\text{“}n\text{ choose }k\text{”: } \binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{(n)(n-1)\cdots(n-k+1)}{(k)(k-1)\cdots(1)}$$

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

$$\sum_{k=0}^{\infty} ar^k = a/(1-r) \qquad \sum_{k=0}^{n} ar^k = (a - ar^{n+1})/(1-r)$$

$$\Pr[X \le x] = F(x) = \int_{-\infty}^{x} f(t)\,dt \qquad f(x) = F'(x)$$

$$\mathsf{E}[g(X)] = \int g(x) f(x)\,dx \qquad \mathsf{E}[g(X)] = \sum_x g(x) P(x)$$

$$\mu = \mathsf{E}[X] = \int_{-\infty}^{\infty} x f(x)\,dx \qquad \sigma^2 = \mathsf{E}[(X-\mu)^2] = \mathsf{E}[X^2] - \mu^2$$

A **and** $B : A \cap B$ **not** $A : A^c$ A **or** $B : A \cup B$

$$\mathsf{P}[A|B] = \mathsf{P}[A \cap B]/\mathsf{P}[B] \qquad \mathsf{P}[B|A] = \mathsf{P}[A \cap B]/\mathsf{P}[A]$$

$$\mathsf{P}[A \cap B] = \mathsf{P}[A|B]\,\mathsf{P}[B] \qquad \mathsf{P}[A \cup B] = \mathsf{P}[A] + \mathsf{P}[B] - \mathsf{P}[A \cap B]$$

Independent $\Leftrightarrow \mathsf{P}[A \cap B] = \mathsf{P}[A]\,\mathsf{P}[B]$; Exclusive $\Leftrightarrow \mathsf{P}[A \cap B] = 0$

Bayes: $\mathsf{P}[A_i|B] = \frac{\mathsf{P}[B|A_i]\mathsf{P}[A_i]}{\mathsf{P}[B|A_1]\mathsf{P}[A_1] + \ldots + \mathsf{P}[B|A_k]\mathsf{P}[A_k]}$ $A_i \cap A_j = \emptyset, \quad \mathsf{P}[A_1] + \ldots + \mathsf{P}[A_k] = 1$

Sterling: $n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$

		Continuous	Discrete
Mean	$\mu = \mathsf{E}[X] =$	$\int_{-\infty}^{\infty} x\, f(x)dx$	$= \sum xp(x)$
Variance	$\sigma^2 = \mathsf{Var}[X] =$	$\int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx$	$= \sum (x-\mu)^2 p(x)$
		$= \mathsf{E}[X^2] - \mathsf{E}[X]^2$	$= \mathsf{E}[X^2] - \mathsf{E}[X]^2$
Expectation	$\mathsf{E}[g(X,Y)] =$	$\iint g(x,y)\, f(x,y)dx\, dy$	$= \sum_{x,y} g(x,y)\, p(x,y)$
Covariance	$\mathsf{Cov}[X,Y] =$	$\iint (x-\mu_x)(y-\mu_y)\, f(x,y)dx\, dy$	$= \sum_{x,y} (x-\mu_x)(y-\mu_y)\, p(x,y)$
		$= \mathsf{E}[XY] - \mathsf{E}[X]\mathsf{E}[Y]$	$= \mathsf{E}[XY] - \mathsf{E}[X]\mathsf{E}[Y]$
Correlation coefficient	$\rho[X,Y] =$	$\mathsf{Cov}[X,Y]/\sigma_X\sigma_Y$	
Conditional	$f(x\|y) =$	$f(x,y)/f_2(y)$	$= p(x,y)/p_2(y)$
Marginal	$f_1(x) =$	$\int_{-\infty}^{\infty} f(x,y)dy$	$= \sum_y p(x,y)$
	$f_2(y) =$	$\int_{-\infty}^{\infty} f(x,y)dx$	$= \sum_x p(x,y)$
Independent		$\Leftrightarrow f(x,y) = f_1(x)\, f_2(y)$	$\Leftrightarrow p(x,y) = p_1(x)\, p_2(y)$
		$\Leftrightarrow f(x\|y) = f_1(x)$	$\Leftrightarrow p(x\|y) = p_1(x)$
		$\Leftrightarrow f(y\|x) = f_2(y)$	$\Leftrightarrow p(y\|x) = p_2(y)$

Conf. Intervals	$\mu = \bar{x} \pm z_{\alpha/2} S/\sqrt{n}$	Large Sample
	$= \bar{x} \pm t_{\alpha/2} S/\sqrt{n}$	Small Sample (normal)
	$p = \hat{p} \pm z_{\alpha/2} \sqrt{\hat{p}\,\hat{q}\,/n}$	Population Proportion
Diff. of Means	$\mu_1 - \mu_2 = (\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	Large Samples
	$= (\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	Small Samples
	$= \overline{x_1 - x_2} \pm z_{\alpha/2} s_\Delta/\sqrt{n}$	Matched Pairs, large n
	$= \overline{x_1 - x_2} \pm t_{\alpha/2} s_\Delta/\sqrt{n}$	Matched Pairs, small n, normal

Sample Mean	$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$
Sample Variance	$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$ Deg fdm: $\nu = n - 1$
Regression slope	$b_1 = \hat{\beta}_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$
Regression offset	$b_0 = \hat{\beta}_0 = \frac{\sum y_i - \hat{\beta}_1 \sum x_i}{n} = \bar{y} - \hat{\beta}_1 \bar{x}$
Sum Square Error	$SSE = \sum (y_i - \hat{y}_i)^2$
Estimate of variance	$\hat{\sigma^2} = \frac{SSE}{n-2}$
TSS	$SST = \sum (y_i - \bar{y})^2$
CD	$r^2 = 1 - \frac{SSE}{SST}$
F-dist	$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$ F-dist with $m-1$,$n-1$ d.o.f.