SIEO 3600: Midterm practice problems. Spring 2008.

This is illustrative of what to expect on the midterm exam: 5 problems. Other topics that might be covered on the exam: moment generating functions, lognormal distribution for stock pricing, Baye's theorem.

1. (20 points) A box has width W, height H and length L (all in centimeters). W and H have a continuous joint distribution over $(1,3)$, given by

$$f(w,h) = \begin{cases} cwh, & 1 < h < w < 3; \\ 0, & \text{otherwise,} \end{cases}$$

where $c > 0$ is a constant. Independently, L has the discrete distribution given by $P(L=4) = 0.25,\ P(L=8) = 0.75$. Let V denote the volume of the box.

(a) (10 points) Find the value of c.

SOLUTION:
A density must integrate out to 1, hence we must solve

$$c\int_1^3\int_1^w wh\ dhdw = 1,$$

equivalently

$$c\int_1^3\int_h^3 wh\ dwdh = 1.$$

Thus

$$c = \Big[\int_1^3\int_1^w wh\ dhdw\Big]^{-1} = 1/8.$$

(Computation left to reader.)

(b) (10 points) Compute $E(V)$

SOLUTION: $E(V) = E(WHL) = E(WH)E(L)$ since L is independent of (W,H). Now, $E(L) = 4(0.25) + 8(0.75) = 7$, while $E(WH) = \int_1^3\int_1^w (wh)f(w,h)\ dhdw = (1/8)\int_1^3\int_1^w w^2h^2\ dhdw = 4.694$. (Note: since both W and H are ≤ 3, and ≥ 1 we know apriori that $1 \leq E(WH) \leq 9$.) Thus $E(V) = (4.694)(7) = 32.86$.

2. (20 points)

Suppose we want a continuous rv X that has density $f(x) = 2x,\ x \in (0,1)$.

(a) (10 points) Show that $X = \sqrt{U}$ works, where U denotes a random variable with a continuous uniform distribution over the interval $(0,1)$; $P(U \leq x) = x,\ x \in (0,1)$.

SOLUTION: With $h(u) = \sqrt{u}$, The density of $X = \sqrt{U} = h(U)$ is given by $f_X(x) = f_U(h^{-1}(x))|J(x)|$, where $J(x) = (h^{-1})'(x)$. Here $f_U(u) = 1,\ u \in (0,1)$ and $h^{-1}(x) = x^2$; $J(x) = 2x$. Thus, indeed, $f_Y(x) = 2x,\ x \in (0,1)$.

(b) (10 points) Show that $X = \max\{U_1, U_2\}$ works, where U_1 and U_2 denote independent uniforms over the interval $(0,1)$.

SOLUTION:

Let $M = \max\{U_1, U_2\}$. We must show that M has density $f_M(x) = 2x,\ x \in (0,1)$. Observe that $\max\{U_1, U_2\} \le x$ if and only if BOTH $U_1 \le x$ and $U_2 \le x$ since if the maximum is $\le x$ then so is the minimum, so both must be $\le x$. Using this fact: $F_M(x) = P(\max\{U_1, U_2\} \le x) = P(U_1 \le x,\ U_2 \le x) = P(U_1 \le x)P(U_2 \le x)$ (from independence of U_1 and U_2)
$= x^2$. Thus $f_M(x) = F_M'(x) = 2x,\ x \in (0,1)$.

NOTE: The density $f(x) = 2x$ on $(0,1)$ is an example of a *beta* distribution, and can be used to model proportions.

3. (10 points)

   2% of wine bottles have defective corks. Compute (approximately) the probability that out of a random sample of 150 bottles, no more than 2 have defective corks.

   SOLUTION: $X =$ the number of bottles out of 150 that are defective has a binomial(n, p) distribution with $n = 150$ trials and success probability $p = 0.02$, and we want $P(X \le 2)$. Here, since $p = 0.02$ is small, while $n = 150$ is large, we can use the Poisson approximation with mean $np = 150(.02) = 3$. $P(X \le 2) = P(X = 0) + P(X = 1) + P(X = 2) \approx e^{-3}(1 + 3 + 3^2/2) = 0.423$

4. The ideal size of a first-year class at a particular college's engineering school is 150 students. The college, knowing from past experience that, on average, only 30% of those accepted for admission will actually attend, uses a policy of accepting 450 students. Compute the probability that more than 150 students attend this college.

   SOLUTION: $X = \sum_1^{450} X_i$, where the X_i are iid Bernoulli(p) rvs with $p = 0.30$, denotes the number of accepted students who will actually attend, has the binomial (n, p) distribution $(n = 450,\ p = 0.30)$ and we want $P(X > 150) = P(X \ge 151)$. Here $n = 450$ is very large so the normal approximation is easily justified (and also p is relatively close to 0.5). With Z denoting a standard $N(0,1)$ rv, we let $Y = \sigma Z + \mu$ where $\mu = np = 450(0.30) = 135$ and $\sigma = \sqrt{np(1-p)} = \sqrt{450(0.30)0.70)} = 9.72$. We use a correction amount of 0.50 since Y is a continuous rv and the binomial is discrete; we use the approximation $P(X \ge 151) \approx P(Y \ge 150.5) = P(Z \ge (150.5 - \mu)/\sigma) = P(Z \ge 1.59) = 1 - P(Z \le 1.59) = 1 - \Phi(1.59) = 1 - 0.9441 = 0.06$.

5. The lifespan of Great Dane dogs has an average of 8 years and a variance of 4. Let A denote the event { a Great Dane lives for more than 12 years or less than 4 years}. (a) At most, what is $P(A)$? (b) Compute $P(A)$ exactly if the lifespan is known to have a continuous uniform distribution over the time interval $(8 - 2\sqrt{3}, 8 + 2\sqrt{3})$ years. (c) Re-do if the distribution of X is discrete as follows: $P(X = 8) = 21/25$, $P(X = 3) = P(X = 13) = 2/25$.

   SOLUTION: (a) Let X denote the lifespan, a rv. All we know is that $E(X) = \mu = 8$ and $\sqrt{Var(X)} = \sigma = 2$. The event $A = \{X > 12 \text{ or } X < 4\}$ can be re-written as $\{|X - 8| > 4\}$, or $\{|X - \mu| > k\sigma\}$ where $k = 2$. Chebychev's inequality yields $P(A) = P(|X - \mu| > k\sigma) \le 1/k^2 = 1/4$, that is, $P(A) \le 1/4$ for ALL distributions on X having mean $\mu = 4$ and standard deviation $\sigma = 2$. (b) In this case X indeed has mean 8 and $\sigma = 2$, but since $2\sqrt{3} = 3.46$, X is bounded between 4.53 and 11.46, and hence can NEVER be larger than 12 nor smaller than 4; $P(A) = 0$. (c) In this case X again has mean 8 and $\sigma = 2$, and $P(A) = P(X = 3) + P(X = 13) = 4/25$. (The point here is that Chebychev's inequality yields a bound, and our first example shows how crude the bound can be.)