**BST 140.651 Midterm Exam**

Notes:

- Show your work on all questions. Simple “yes” or “no” answers will be graded as if blank.
- Please be neat and write legibly. Use the back of the pages if necessary.
- There are eight questions.
- Good luck!

____________________

signature and **printed name**

1. Yearly medical expenditures (in $1000$ dollars) for a particular demographic follow the density $f(x) = 2x\exp(-x^2)$ for $x > 0$. Show mathematically that this is a valid density.

Note $x > 0$ so that

$2x > 0$ & $e^{-x^2}$ is always $> 0$

so that $2x e^{-x^2} \geq 0$

Secondly $\displaystyle\int_0^\infty 2xe^{-x^2}dx = -e^{-x^2}\Big|_{x=0}^{\infty} = 0-(-1)$

$= 1$

2. Refer to the previous problem. What is the median level of expenditure in this population? (Note, do not solve for a final number, just carry the mathematics until the point of solving for a final number.)

We want the x so that

$$\int_0^x 2t e^{-t^2} dt = \int_x^\infty 2t e^{-t^2} dt = .5$$

Note

$$\int_x^\infty 2t e^{-t^2} dt = -e^{-t^2}\Big|_{t=x}^{\infty} = 0 - (-e^{-x^2})$$
$$= e^{-x^2}$$

Setting $= .5$ we get

$$.5 = e^{-x^2} \quad *$$

$$\Rightarrow -\log(2) = -x^2 \quad \Rightarrow \quad x = \sqrt{\log(2)}$$

Note if you do it the other way

$$\int_0^x 2t e^{-t^2} dt = -e^{-t^2}\Big|_0^x = -e^{-x^2} + 1 = .5$$

$$\Rightarrow \quad e^{-x^2} = .5 \quad \text{same as } * \text{ above}$$

3. Refer to the two previous problems. The **mode** of a density is the point at which density is the highest; that is, the point $x_m$ so that $f(x_m) > f(x)$ for any other $x$. Calculate the modal medical expenditure for this population. (Hint, take a log before differentiating.)

$$\log\{f(x)\} = \log(2) + \log(x) - x^2$$

$$\frac{d}{dx}\log\{f(x)\} = \frac{1}{x} - 2x \quad \text{set} = 0$$

we obtain $\quad 2x = 1/x \quad \Rightarrow \quad x = \sqrt{\frac{1}{2}}$

2nd deriv condition

$$\frac{d^2}{dx^2}\log\{f(x)\} = \underbrace{-\frac{1}{x^2}}_{<0} \underbrace{-2}_{<0}$$

so the function is concave

4. Let $X_1$ and $X_2$ be iid random variables from a population with mean $\mu_1$ and variance $\sigma_1^2$ and $Y_1$ and $Y_2$ be iid random variables from a population with mean $\mu_2$ and variance $\sigma_2^2$. What is the **expected value** of $(X_1 + X_2)^2 + (Y_1 + Y_2)^2$?

$$E\left[(x_1+x_2)^2 + (y_1+y_2)^2\right] =$$

$$E\left[(x_1+x_2)^2\right] + E\left[(y_1+y_2)^2\right] =$$

$$Var(x_1+x_2) + \left(E[x_1]+E[x_2]\right)^2 + Var(y_1+y_2) + \left(E(y_1+y_2)\right)^2$$

$$= 2\sigma_1^2 + 4\mu_1^2 + 2\sigma_2^2 + 4\mu_2^2$$

Alternatively

$$E\left[(x_1+x_2)^2\right] + E\left[(y_1+y_2)^2\right]$$

$$= E\left[4\bar{x}^2\right] + E\left[4\bar{y}^2\right] \qquad \bar{x} = \frac{x_1+x_2}{2}$$

$$= 4\left(Var(\bar{x}) + E[\bar{x}]^2\right) + 4\left(Var(\bar{y}) + E(\bar{y})^2\right) \qquad \bar{y} = \frac{y_1+y_2}{2}$$

$$= 4\left(\frac{\sigma_1^2}{2} + \mu_1^2\right) + 4\left(\sigma_2^2/2 + \mu_2^2\right)$$

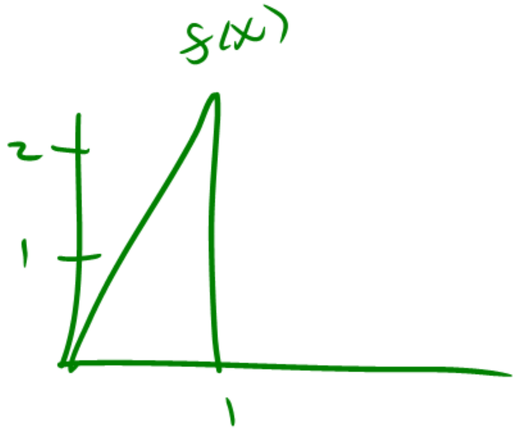
$$= 2\sigma_1^2 + 4\mu_1^2 + 2\sigma_2^2 + 4\mu_2^2$$

5. Let $X$ be a random variable, draw the density $f(x) = 2x$ for $0 \le x \le 1$. Calculate $E[X]$ and $E[X^2]$.

$$E[X] = \int_0^1 x \, 2x \, dx = 2\int_0^1 x^2 \, dx = \frac{2}{3} x^3 \Big|_0^1 = \frac{2}{3}$$

$$E[X^2] = \int_0^1 x^2 \, 2x \, dx = 2 \int_0^1 x^3 \, dx = \frac{2}{4} x^4 \Big|_0^1 = \frac{1}{2}$$

density



6. An HIV antibody test is known to be 90% sensitive and 80% specific. By what factor are the prior odds of disease increased with a positive test result? By what fraction are the prior odds of disease decreased with a negative test result?

$$DLR_+ = \frac{Sens}{(1 - Spec)} = \frac{.90}{.20} = 4.5$$

$$DLR_- = \frac{1 - Sens}{Spec} = \frac{.1}{.8} = \frac{1}{8}$$

7. Three events, $A$, $B$ and $C$ all have the same probability of occurrence; call this probability $p$. Argue that the probability of at least one occurring is less than or equal to $3p$. (You may assume any result proved in class or in the homework.)

Note we showed in class that

$$P(A \cup B \cup C) \leq P(A) + P(B) + P(C) = 3p$$

Alternatively

$$P(A \cup B \cup C) = P(A) + P(B \cup C) \underbrace{- P(A \cap (B \cup C))}_{\text{since } \leq 0}$$

$$\leq P(A) + P(B \cup C)$$

$$= P(A) + P(B) + P(C) \underbrace{- P(B \cap C)}_{\leq 0}$$

$$\leq P(A) + P(B) + P(C)$$

$$= p + p + p = 3p$$

8. Consider a diagnostic test for tumors that yields two possible diagnoses, $T$ for tumor and $T^c$ for no tumor. Potential tumors for this cancer are always one of: malignant ($D_M$), benign ($D_B$) or not tumors ($D_n$). Argue the generalized version of Bayes rule:

$$P(D_M \mid T) = \frac{P(T \mid D_M)P(D_M)}{P(T \mid D_M)P(D_M) + P(T \mid D_B)P(D_B) + P(T \mid D_N)P(D_N)}.$$

You may use any result proved in class or in the homework.

Look at the RHS

$$\frac{P(T \mid D_M)\,P(D_M)}{P(T \mid D_M)P(D_M) + P(T \mid D_B)P(D_B) + P(T \mid D_N)P(D_N)}$$

$$= \frac{P(T \cap D_M)}{P(T \cap D_M) + P(T \cap D_B) + P(T \cap D_N)}$$

$$= \frac{P(T \cap D_M)}{P(T)} = P(D_M \mid T)$$