MIDTERM EXAM

(1) (35 points) Suppose $\{X_i\}_{i=1}^{9}$ are an i.i.d. sequence of $N(\mu,1)$ random variables.

(a) (5 points) Show that the sample mean of the X_i is a sufficient statistic for μ.

The pdf of the sample is

$$(2\pi\sigma^2)^{-4.5}e^{-\frac{1}{2}\Sigma(x_i-\mu)^2} = (2\pi\sigma^2)^{-4.5}e^{-\frac{1}{2}\Sigma(x_i-\bar{x})^2}e^{-4.5(\bar{x}-\mu)^2}\,.$$

Only the last exponential term in the right-hand side above depends on μ, which by the factorization criterion tells us that $\bar{x}$, which is the only function of x entering that term, is a sufficient statistic.

(b) (5 points) Use the fact that the cdf $\Phi(z)$ for a $N(0,1)$ random variable satisfies $\Phi(1.96) = .975$ to form a 95% confidence interval for μ when the sample mean of these 9 observations is $\bar{X} = .1$.

The sample mean is itself normal, with mean μ and variance $1/9$. Also, the normal distribution is symmetric about zero. So for every μ,

$$P\left[\bar{X} \in \left(\frac{-1.96}{3}+\mu, \frac{1.96}{3}+\mu\right) \middle| \mu\right] = .05\,.$$

Since the event whose probability is described here can also be written as

$$\mu \in (\bar{X} - .653, \bar{X} + .653))\,,$$

The random interval on the right in this expression is a 95% confidence interval for μ, and with our observed $\bar{X} = .1$, this is $(-.553, .753)$.

(c) (5 points) Labeling the ends of your confidence interval from 1b as (a,b), it might seem that if (a,b) is a 95% confidence interval for μ, $(1/b, 1/a)$ would be a valid 95% confidence interval for μ^{-1}. Is this correct? If so, explain why. If not, explain why not and present a valid confidence interval for μ^{-1} for the same $\bar{X} = .1$ case considered in part 1b.

The interval suggested would be appropriate if a and b were always positive, but this is not true here, and in our sample the proposed interval would be empty (having left endpoint larger than the right endpoint). Mapping the interval for μ into an interval for $1/\mu$ will provide a 95% confidence set, but here that mapping (since the original interval includes zero, which maps into ∞) gives us the confidence set $(-\infty, -1/.553) \cup (1/.753, \infty)$. Note that this is not an "interval", so does not quite respond to the question's request for a valid confidence interval. It is an interesting question whether there *is any straightforward way to construct a valid 95% confidence set for $1/\mu$ that is a connected interval with probability one.*

(d) (5 points) Find the posterior pdf for μ with this same sample, assuming as prior $\mu \sim N(0, 2500)$. Use it to construct a 95% minimum-length posterior probability interval for μ.

Prior pdf times likelihood is proportional to (dropping constants)

$$exp\left(-\frac{1}{2}\left(\sum(x_i-\mu)^2+\frac{\mu^2}{2500}\right)\right)=exp\left(-\frac{1}{2}\left((9+.0004)\mu^2-2\cdot 9\mu\bar{x}+\sum(x_i^2)\right)\right).$$

Completing the square in μ in this expression gives us $9.0004\cdot(\mu-(9/9.0004)\bar{x})^2$ plus terms not dependent on μ. To two-significant-figure accuracy, this is just $9\cdot(\mu-\bar{\mu})^2$. Thus the posterior is proportional to a $N(\bar{x}, 1/9)$ pdf and a minimum-length 95% probability interval for μ is the same as the 95% confidence interval we found above.

(e) (10 points) Find the posterior pdf for $\nu = \mu^{-1}$, still assuming the same sample and assuming that the prior is still $\mu \sim N(0, 2500)$. [Of course you will have to convert the prior on μ to one on ν.] Sketch the shape of this pdf, showing the rough location of its maximum (or maxima, if there are multiple local maxima).

We convert the posterior for μ into a posterior for $1/\mu$ using the Jacobian rule, which tells us that the posterior pdf is $\varphi(\mu-\bar{x}; 1/9)\cdot\mu^2$. This clearly has a zero at $\mu = 0$ and declines to zero eventually at $\pm\infty$ and therefore must have at least two peaks, one on either side of zero. Proving that there are exactly two peaks requires some further argument. For example, one can look at the derivative of the posterior pdf, which is proportional to

$$(2\mu-9\mu^2(\mu-\bar{x}))e^{-\frac{9}{2}(\mu-\bar{x})^2}.$$

since the exponential part of this is always positive, the zeros are the zeros of the first factor, which as a 3rd order polynomial has no more than three zeros. Therefore we have just the two local maxima, plus the local minimum at $\mu = 0$. The higher of the two peaks will be positive when $\bar{x} > 0$, and vice versa when $\bar{x} < 0$.

(f) (5 points) Describe and/or sketch the shape of a minimum-length 95% posterior probability interval for ν.

A minimum-length "interval" will have pdf values equal at all its boundary points. Here that means the confidence set will have probably have two pieces, one shorter, to the left of zero, and one longer, to the right of zero. It is of course possible in general, when $\bar{x}$ is far enough from zero and the smaller peak is therefore low enough, for

the set to consist of a single interval. Here, because $\bar{x}$ is only about one standard deviation away from zero, one expects that the interval will consist of two pieces.
Verifying that carefully was more than was expected on the exam. One approach to doing so runs as follows. Since $P[\mu > 0] = P[1/\mu > 0]$, This probability is, for our $N(.1, .1111)$ posterior, $\Phi(.9)$, i.e. the probability that a standard normal variate is less than .9 standard deviations from zero. But $\Phi(1) = .84$, approximately, so no interval entirely to the right of zero can have 95% probability.

*[In items 1e-1f you won't be able to complete exact numerical calculations. For the other parts of this question numerical results are probably possible with pencil and paper. Accuracy to two significant figures suffices, so in **some** cases you may find very dispersed proper priors give the same numerical answers as flat priors.]*

(2) (15 points) For each of the following functions, state whether or not it is a cdf. If it is a cdf, state whether or not the two random variables it defines are independent and whether or not they have a pdf with respect to Lebesgue measure. Explain your answers.

(a) (5 points)

$$F(x,y) = \begin{cases} 1 - e^{-x} - e^{-y} + e^{-x-y} & x,y \text{ both positive} \\ 0 & x < 0 \text{ or } y < 0 \end{cases}$$

This function can be written $(1-e^{-x})(1-e^{-y})$, which is the product of two univariate exponential cdf's. It is therefore itself a cdf, has a pdf w.r.t. Lebesgue measure, and makes X and Y independent.

(b) (5 points)

$$F(x,y) = \begin{cases} \sqrt{x^2+y^2} & 0 \le x^2+y^2 < 1\,,\, x \ge 0\,,\, y \ge 0 \\ 1 & x^2+y^2 \ge 1 \\ 0 & x < 0 \text{ or } y < 0 \end{cases}$$

While this function is monotone increasing, and lies between 0 and 1, it is not a cdf. A quick way to see this is to observe that $F(x,\infty) = F(\infty,y) = 1$ for x or y positive, and zero otherwise. Thus the marginal distributions imply all probability is concentrated at $x = y = 0$, but a cdf for that joint distribution would $F(x,y) = 1$, whenever x and y were both positive. Another way to see the point is to observe that the function implies the probability of the unit square is $F(1,1) + F(0,0) - F(1,0) - F(0,1) = -1$.

(c) (5 points)

$$F(x,y)=\begin{cases} 2\big(\arcsin(y)-\arccos(x)\big)/\pi & \arcsin(y)>\arccos(x)\,,\, x\in(0,1)\,,\, y\in(0,1)\\ 2\arcsin(y)/\pi & y\in(0,1)\,,\, x>1\\ 1-2\arccos(x)/\pi & x\in(0,1)\,,\, y>1\\ 1 & x>1\,,\, y>1\\ 0 & \text{otherwise}\end{cases}$$

Here $\arcsin(x)$ is the inverse sine function, defined by $\sin(\arcsin(x))=x$, and similarly for arccos.

This is a cdf, does not have a density w.r.t. Lebesgue measure, and does not make X and Y independent. That it does not have a density, assuming for now that it is indeed a cdf, can be seen from the fact that $\partial^2 F/\partial x\partial y=0$ everywhere except at the boundaries of the regions of definition for F. So F is not the integral of its of its cross partial derivative, which would be its density if it had one.

The cdf in fact concentrates all probability on the unit quarter-circle in the positive quadrant. To see this, note that $\arccos(x)=\arcsin(\sqrt{1-x^2})$, so that, for x and y both in $(0,1)$, The condition $\arcsin(y)>\arccos(x)$ is equivalent to $x^2+y^2>1$. So the cdf is zero inside the unit circle and increases as the point x,y moves outside it, reaching its max when $x=y=1$.

To prove that it is in fact a cdf one could look at implied probabilities of rectangles, showing that for any rectangle with corners (a,b) and (c,d), with $a>c$, $b>d$, the probability, which is given by $F(a,b)+F(c,d)-F(a,d)-F(c,b)$, is non-negative. Because the function is differentiable outside the unit square and has zero cross partial there, we know the probability of a rectangle with (a,b) outside the square is the same as the probability of the intersection of that rectangle with the unit square, so we need only consider cases where (a,b) and (c,d) are both inside the unit square. When both endpoints are inside the unit circle, or both outside, it is easy to see that the probability of the rectangle is zero. When (a,b) is outside and (c,d) inside, there are two cases. One is that (a,d) and (c,b) are both inside. Then the rectangle's implied probability is just $F(a,b)\geq 0$. When, say, (a,d) is outside and (c,b) is inside, the probability is $\arccos b-\arccos d>0$, and when (a,d) is inside while (c,b) is outside, the probability is $\arcsin a-\arcsin c>0$.

(3) (10 points) Suppose $\{X_i\}_{i=1}^n$ is an i.i.d. sequence and we wish to test $H_0: X_i\sim N(0,1)$ against the alternative $H_A:\ X_i\sim N(0,\sigma^2)$, $\sigma^2>1$. We begin by forming the test statistic $V=\bar{X}/\sqrt{S_x^2}$, where $\bar{X}$ is the sample mean and $S_X{}^2=\sum(X_i-\bar{X})^2/(n-1)$.

It is known that on the null hypothesis this statistic has a t_{n-1} distribution. We form a rejection region by looking up in the t table a value such that $P[t_{n-1} < \theta] = .975$, then setting the rejection region to be $\{V > \theta\} \cup \{V < -\theta\}$.

(a) (2 points) Is this a test with size (i.e. significance level) .05?

Yes, because the symmetry of the t distribution makes its rejection region have probability $2 \cdot .025$ *under* H_0.

(b) (4 points) Is it a biased test?

A biased test has the property that for some point in H_A *the probability of rejection is lower than the probability of rejection under some point in* H_0. *So we must figure out what is the probability of rejection under* H_A. *But notice that under* H_A, $\bar{X}/\sigma$ *has the same distribution as* $\bar{X}$ *under* H_0, *while* S_X^2/σ^2 *has the same distribution as* S_X^2 *under* H_0. *The ratio* $\bar{X}/\sqrt{S_X^2}$ *therefore has the same distribution under the null and the alternative. The rejection probability is therefore the same at every point in the parameter space and the test is unbiased.*

(c) (4 points) Explain why tests with better power are surely available.

This test is completely uninformative — we might as well draw marbles from a hat with 19 white and 1 black marble, and reject when we get the black marble. This nonsense test would have the same significance level and power as the t test proposed here.
In this model the likelihood function is proportional to

$$\sigma^{-n} \exp\left(-\tfrac{1}{2}\sum_{1}^{n}\frac{x_i^2}{\sigma^2}\right).$$

A sufficient statistic for the one unknown parameter, σ^2, *is therefore* $\sum x_i^2$. *It is clear that a better test could be constructed based on this statistic alone. This completes a perfect answer.*
If you knew or remembered that the sum of n i.i.d. $N(0,1)$ *random variables is distributed as* $\chi^2(n)$, *you might have proposed the test that is in fact best here by many criteria: reject when* $\sum x_i^2$ *exceeds the .95 level of the* $\chi^2(n)$ *distribution.*

(4) (15 points) The random variables in the sequence $\{X_i\}_{i=1}^{\infty}$ are distributed with $\Gamma(p_i, \alpha_i)$ distributions, that is they are positive with probability 1 and have pdf's

$$\Gamma(p_i)^{-1}\alpha_i^{p} x_i^{(p_i-1)} e^{-\alpha_i x_i}.$$

The sequence of p's and α's satisfies $p_i = \alpha_i = i$. Show that $P[X_i < 1 - i^{-1/3}] \xrightarrow[i\to\infty]{} 0$.

A typo in this question omitted the minus sign in front of the $1/3$*, so that it became trivial. Since no one noticed this while I was in the room asking about typos, it was legitimate, if not very sporting, to simply note that the probability in question is exactly zero as soon as* $i > 1$*. What I give here is the answer to the question as it was meant to be.*

Since the mean of a $\Gamma(p,\alpha)$ *pdf is* p/α*, all these* X_i*'s have the same mean, namely 1. The variance of a* $\Gamma(p,\alpha)$ *is* p/α^2*, so* $\mathrm{Var}(X_i) = 1/i$*. The Chebyshev inequality therefore tells us*

$$P[X_i - 1 < -i^{-1/3}] < \mathrm{Var}(X_i) i^{2/3} = i^{-1/3} \xrightarrow[i\to\infty]{} 0\,.$$

(5) (15 points) The two-dimensional random vector X satisfies

$$X \sim N\left(\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}9 & 2\\2 & 6\end{bmatrix}\right)\,.$$

Find the maximum of the pdf and the vectors that are the directions of the major and minor axes of the ellipses that form its level curves. Sketch the level curves of the pdf, marking on your plot the location of the peak of the pdf and the major and minor axes of the level curves.

The maximum of a normal pdf occurs at its mean, here the point $[1,2]'$*. The major and minor axes of the level curves are in the directions of the eigenvectors corresponding to the largest and smallest elements, respectively, of the covariance matrix. Solving*

$$\begin{bmatrix}9 & 2\\2 & 6\end{bmatrix}\begin{bmatrix}1\\a\end{bmatrix} = \begin{bmatrix}\lambda\\\lambda a\end{bmatrix}$$

gives us $a = -2$ *and* $a = .5$ *as solutions. The* $[1,-2]$ *vector corresponds to an eigenvalue of 5, while the* $[1,.5]$ *vector corresponds to an eigenvalue of 10. The level curves therefore look like the graph below.*