“Core Competency” (**Fall 2018**) **Name (Print):** ____________________
Qualifying Exam (09/04/2018)
Time Limit: 4 hours **Signature:** ____________________

This exam contains 9 problems. Answer **all** of them. Point values are in parentheses. You **must show your work** to get credit for your solutions — correct answers without work will not be awarded points.

No calculator will be allowed in the exam. This is a closed notes/book exam; no cheat sheets are allowed. Normal tables, if needed, will be provided during the exam. Rough paper will be provided to you.

1	7 pts	
2	8 pts	
3	10 pts	
4	15 pts	
5	15 pts	
6	15 pts	
7	8 pts	
8	10 pts	
9	12 pts	
TOTAL	100 pts	

1. (7 points) Consider a routine screening test for a disease. Suppose the frequency of the disease in the population (base rate) is 0.5%. The test is highly accurate with a 5% false positive rate and a 10% false negative rate [a false positive happens when a test result indicates that the disease is present (the result is positive), but it is, in fact, not present. Similarly, a false negative happens when a test result indicates that the disease is not present (the result is negative), but it is, in fact, present].

   You take the test and it comes back positive. What is the probability that you have the disease?

2. (8 points) (4+4) We consider balls of random radius R.

   (i) Suppose that R is uniformly distributed on $[1, 10]$. Find the probability density function of the volume V of a ball. (Recall that $V = \frac{4}{3}\pi R^3$.)

   (ii) Suppose that R has a log-normal distribution, meaning that $\log(R) \sim \mathcal{N}(\mu, \sigma^2)$ for some parameters $\mu \in \mathbb{R}$ and $\sigma > 0$. Show that V also has a log-normal distribution and find its parameters.

3. (10 points) (6 + 4) Suppose that, for $n \geq 1$, X_n is a random variable taking values in $\{\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}\}$ with equal probability $\frac{1}{n}$.

   (i) Show that X_n converge in distribution, as $n \to \infty$? What is its weak limit?

   (ii) Let $f : [0,1] \to \mathbb{R}$ be defined as $f(x) = x\sin(x)$, for $x \in [0,1]$. Using the above or otherwise, show that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f\left(\frac{k}{n}\right) = \int_0^1 f(x)\, dx.$$

4. (15 points) (5 + 4 + 6)

(i) Let X be a random variable and $a \in \mathbb{R}$. Show that (using Markov's inequality or otherwise):

$$\mathbb{P}[X \geq a] \leq \inf_{s \geq 0} e^{-sa}\, \mathbb{E}[e^{sX}].$$

(ii) Let N be a Poisson random variable with parameter $\lambda > 0$; i.e.,

$$\mathbb{P}[N = n] = e^{-\lambda} \frac{\lambda^n}{n}, \quad n \geq 0.$$

Show that $\mathbb{E}[e^{sN}] = e^{\lambda(e^s - 1)}$ for all $s \in \mathbb{R}$.

(iii) Let N be as in (ii) and let $m \geq \lambda$ be an integer. Use (i) and (ii) to show that

$$\mathbb{P}[N \geq m] \leq \left(\frac{\lambda}{m}\right)^m e^{m-\lambda}.$$

5. (15 points) (3 × 5) We obtain observations $Y_1, \ldots, Y_n$ which can be described by the relationship

$$Y_i = i \times \theta + \epsilon_i,$$

where $\epsilon_1, \ldots, \epsilon_n$ are i.i.d $N(0, \sigma^2)$; $\sigma^2 > 0$. Here θ and σ^2 are unknown.

(i) Find the least squares estimator $\hat{\theta}$ of θ; i.e., $\hat{\theta} = \arg\min_{\theta \in \mathbb{R}} \sum_{i=1}^{n} (Y_i - i\theta)^2$.

(ii) Is $\hat{\theta}$ unbiased?

(iii) Find the exact distribution of $\hat{\theta}$.

(iv) Find the asymptotic (non-degenerate) distribution of $\hat{\theta}$ (properly normalized).

(v) How would you test the hypothesis $H_0 : \theta = 0$ versus $H_1 : \theta \neq 0$ (at level $\alpha \in (0, 1)$)? Describe the test statistic and the critical value.

6. (15 points) (3 + 5 + 3 + 4) Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d $N(\theta, 1)$, where $\theta \in \mathbb{R}$ is unknown. Let $\psi = \mathbb{P}_\theta(X_1 > 0)$.

   (a) Find the maximum likelihood estimator $\hat{\psi}$ of ψ.

   (b) Find an approximate 95% confidence interval for ψ.

(c) Let $Y_i = \mathbf{1}\{X_i > 0\}$, for $i = 1, \ldots, n$. Define $\tilde{\psi} = (1/n)\sum_{i=1}^{n} Y_i$. Show that $\tilde{\psi}$ is a consistent estimator of ψ.

(d) Find the asymptotic distribution of both the estimators? Which estimator of ψ, $\hat{\psi}$ or $\tilde{\psi}$, is more preferable in this model and why?

7. (8 points) (3+5) Suppose $X_1, \ldots, X_n$ are i.i.d. with $\mathbb{P}(X_i = \pm 1) = \frac{1}{2}$. Define

$$Y_i := \prod_{j=1}^{i} X_j, \qquad \text{for } i = 1, \ldots, n.$$

(i) Find the joint distribution of (Y_1, Y_2).

(ii) Derive the limiting distribution of $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i$.

8. (10 points) (4+6) Suppose $(\mathbf{X}, \mathbf{Y})$ have a multivariate normal distribution with mean vector $\mathbf{0}$ and covariance matrix

$$\Sigma = \begin{bmatrix} A & B \\ B^\top & C \end{bmatrix},$$

where A is $m \times m$, B is $m \times n$, and C is $n \times n$, and A and C are non-singular. Define a vector $\mathbf{Z} := \mathbf{Y} - B^\top A^{-1}\mathbf{X}$.

(i) Find the $m \times n$ covariance matrix of $\mathbf{X}$ and $\mathbf{Z}$.

(ii) Express $\mathbf{Y}$ as $\mathbf{Z} + B^\top A^{-1}\mathbf{X}$, and hence deduce the conditional distribution of $\mathbf{Y}$ given $\mathbf{X} = \mathbf{x}$.

9. (12 points) (6+6) Let $X \in \mathbb{R}^d$ be a centered normal random vectors and $A \in \mathbb{R}^{d\times d}$ a fixed symmetric matrix.

   (i) Denote by Y an independent copy of X. Show that

   $$X^\top AX - Y^\top AY =_d 2X^\top AY.$$

   Hint: $(X \pm Y)/\sqrt{2}$ are i.i.d. random vectors following the same distribution as X.

(ii) Show that for any $t \geq 0$,

$$\mathbb{P}(|X^\top AX - \mathbb{E}(X^\top AX)| \geq t) \leq \mathbb{P}(|X^\top AY| \geq t).$$

Hint: First show that for two i.i.d. random variables Z_1, Z_2, $\mathbb{P}(|Z_1 - Z_2| \geq 2t) \geq P(|Z_1| \geq t)$.