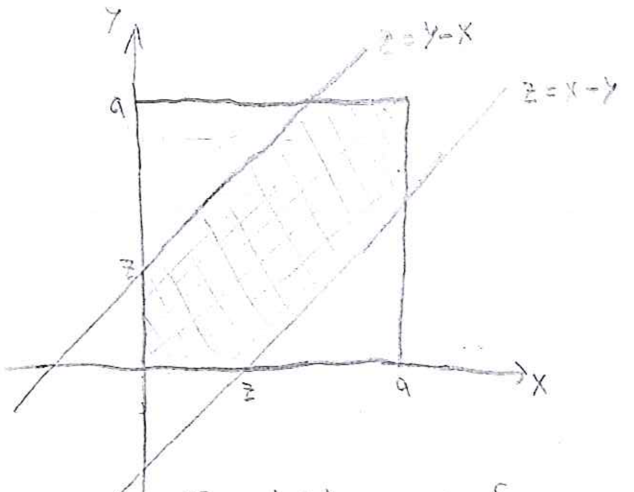
Use this space to show your work for Problem 1 only.



The shaded region is S

$$F_Z(z) = P(Z \le z) = P(|X-Y| \le z)$$

$$= P(-z \le X-Y \le z)$$

$$= \int_S \frac{1}{a^2}\, dS = \frac{1}{a^2}\int_S dS = \frac{1}{a^2} \times (\text{Area of } S)$$

$$= \frac{1}{a^2} \cdot 2\left(\frac{a^2}{2} - \frac{(a-z)^2}{2}\right) = \frac{1}{a^2}\left(a^2 - (a-z)^2\right)$$

$$= \frac{1}{a^2}\left(2az - z^2\right) = \frac{2z}{a} - \frac{z^2}{a^2} \quad \text{when } 0 \le z \le a$$

when $z < 0$, $P(Z \le z) = 0$

when $z > a$, $P(Z \le z) = 1$

$$\Rightarrow F_Z(z) = \begin{cases} 0, & z < 0 \\ \frac{2z}{a} - \frac{z^2}{a^2}, & 0 \le z \le a \\ 1, & \text{otherwise} \end{cases}$$

P1. (5 points)

Let X and Y be two independent drawings from the uniform distribution on $[0, a]$, with $a > 0$ a given constant. Let $Z = |X - Y|$ be the distance between the two points. Find the CDF $F_Z(z)$. Show your work by drawing a figure that explains how you calculate $F_Z(z)$.

$$F_Z(z) = \begin{cases} 0 & , \quad z < 0 \\ \frac{2z}{a} - \frac{z^2}{a^2} & , \quad 0 \leq z \leq a \\ 1 & , \quad z > a \end{cases}$$

Use this space to show your work for Problem 2 only.

$$X_i = \begin{cases} 1, & \text{if the ith digit drawn is } 7. \\ 0, & \text{otherwise.} \end{cases}$$

$\Rightarrow Y = X_1 + X_2 + X_3 + \dots + X_{10000}$

$\Rightarrow P(Y < 968) = P(X_1 + X_2 + X_3 + \dots + X_{10000} < 968)$

$\Rightarrow$ mean of $X_1 + X_2 + X_3 + \dots + X_{10000} = 10000 * (\text{mean of } X_i) = 10.000 \times \frac{1}{10} = 1000 = E[Y]$

$\Rightarrow$ variance of $X_1 + X_2 + X_3 + \dots + X_{10000} = 10000 * \text{var}(X_i) = 10000 \times \left(\frac{1}{10} \cdot \frac{9}{10}\right) = 900 = \text{var}(Y)$

$\Rightarrow$ By CLT,

$$P(Y < 968) = P\left(\frac{X_1 + X_2 + X_3 + X_4 + \dots + X_{10000} - 1000}{\sqrt{900}} < \frac{968 - 1000}{\sqrt{900}}\right) = \Phi\left(\frac{968-1000}{30}\right) = \Phi\left(\frac{-16}{15}\right)$$

$$= 1 - \Phi\left(\frac{16}{15}\right) = 1 - \Phi(1.0\bar{6}) \simeq 1 - \Phi(1.07) = 1 - 0.8577 = 0.1423$$

Note: We could use $\Phi\left(\frac{967.5-1000}{30}\right)$ $\left(\text{or even } \Phi\left(\frac{967.001-1000}{30}\right), \text{etc.}\right)$ for a better estimation, as the case is discrete.

P2. (5 points)

Let Y be the number of times the digit 7 appears among $n = 10,000$ digits drawn from the set $\{0,1,2,3,4,5,6,7,8,9\}$ independently at random (using the uniform distribution). Use the Central Limit Theorem (CLT) to approximate the probability $P(Y < 968)$ in terms of the function $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du$. Use the table given below to give a numerical estimate of probability. **(Show clearly the expectation and variance of Y. Show how you standardize the Y to obtain the CLT approximation. An answer without sufficient justification will receive no credit even if correct.)**

$P(Y < 968) \approx$ 0.1423

Table 1: Standard Normal Table $\Phi(x)$.

	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Use this space to show your work for Problem 3 only.

---

a) $H_0: \Theta = 0$, $H_1: \Theta = 1$

$$p_\Theta(0)\cdot p_{Y|\Theta}(y|0) \underset{H_1}{\overset{H_0}{\gtrless}} p_\Theta(1)\cdot p_{Y|\Theta}(y|1)$$

$$\frac{1}{2}\cdot\frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(y-A)^2/2\sigma^2} \underset{H_1}{\overset{H_0}{\gtrless}} \frac{1}{2}\cdot\frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(y+A)^2/2\sigma^2}$$

$$\Rightarrow e^{-(y-A)^2/2\sigma^2} \underset{H_1}{\overset{H_0}{\gtrless}} e^{-(y+A)^2/2\sigma^2}$$

Decision rule $= \begin{cases} \Theta = 0 & \text{if } Y > 0 \\ \Theta = 1 & \text{if } Y < 0 \end{cases}$

$$\Rightarrow -(y-A)^2 \underset{H_1}{\overset{H_0}{\gtrless}} -(y+A)^2$$

$$\Rightarrow y^2 - 2yA + A^2 \underset{H_1}{\overset{H_0}{\lessgtr}} y^2 + 2yA + A^2$$

$$\Rightarrow -y \underset{H_1}{\overset{H_0}{\lessgtr}} y \quad \Rightarrow \quad 0 \underset{H_1}{\overset{H_0}{\lessgtr}} y$$

$P_e = p_\Theta(0)\cdot P(Y<0 \mid \Theta=0) + p_\Theta(1)\, P(Y>0 \mid \Theta=1)$

$= \frac{1}{2} P(Z < -A) + \frac{1}{2} P(Z > A)$

$= \frac{1}{2} P(Z < -A) + \frac{1}{2}\left(1 - P(Z < A)\right)$ $\quad \boxed{P(Z<-A) = 1-P(Z<A) \text{ as } Z \text{ is normal with } \mu = (}$

$= 1 - P(Z < A)$. With $A = 4$, $\sigma^2 = 4$, using standardization:

$P_e = 1 - P(Z<4) = 1 - P\left(\frac{Z}{2} < 2\right) = 1 - \Phi(2) = 0.0228$

---

b) $\hat{\Theta} = E[\Theta] + \dfrac{\text{cov}(\Theta, X)}{\text{var}(X)}(X - E[X])$, $\quad E[\Theta] = \dfrac{4+10}{2} = 7 \quad E[X] = E[\Theta + W] = E[\Theta] + E[W] = \dfrac{4+10}{2} + \dfrac{1-1}{2} = 7$

$\text{var}(X) = \text{var}(\Theta + W) = \text{var}(\Theta) + \text{var}(W)$ as Θ and W are independent.

$\Rightarrow \text{var}(X) = \dfrac{(10-4)^2}{12} + \dfrac{(1-(-1))^2}{12} = 3 + \dfrac{4}{12} = \dfrac{10}{3}$ $\quad$ (For future use, note $\text{var}(\Theta) = \dfrac{(10-4)^2}{12} = 3$)

$\text{cov}(\Theta, X) = E[\Theta X] - E[\Theta]E[X] = E[\Theta(\Theta + W)] - 49 = E[\Theta^2] + E[\Theta W] - 49$

$= E[\Theta^2] + E[\Theta]E[W] - 49$ (as Θ and W are independent)

$= E[\Theta^2] - 49$ (as $E[W] = \frac{1-1}{2} = 0$)

$E[\Theta^2] = \displaystyle\int_4^{10} \theta'^2 f_\Theta(\theta')\,d\theta' = \int_4^{10} \theta'^2 \frac{1}{6}\,d\theta' = \frac{\theta'^3}{18}\Big|_{\theta'=4}^{10} = \frac{936}{18} = 52$

$\Rightarrow \text{cov}(\Theta, X) = 52 - 49 = 3$

$\Rightarrow \hat{\Theta} = 7 + \dfrac{9}{10}(X - 7) = \dfrac{9X}{10} + \dfrac{7}{10} = \dfrac{9X+7}{10}$

MSE of LLMS $= (1-\rho^2)\sigma_\Theta^2 = \left(1 - \left(\dfrac{\text{cov}(\Theta,X)}{\sigma_\Theta \sigma_X}\right)^2\right)\sigma_\Theta^2 = \left(1 - \left(\dfrac{3^2}{3\cdot\frac{10}{3}}\right)\right)3 = \left(1 - \dfrac{3}{10}\right)3 = \dfrac{3}{10} = 0.3$

P3. (10 points) The two parts of this problem are independent.

(a) (5 pts) Consider a binary communication system in which the message is modeled as an unknown parameter Θ that takes values as

$$p_\Theta(\theta) = \begin{cases} \frac{1}{2}, & \theta = 0; \\ \frac{1}{2}, & \theta = 1. \end{cases}$$

The transmitted signal X is a function of the message Θ:

$$X = \begin{cases} +A, & \text{if } \Theta = 0; \\ -A, & \text{if } \Theta = 1, \end{cases}$$

where $A > 0$ is a constant. The received signal is given by

$$Y = X + Z$$

where $Z \sim N(0, \sigma^2)$ represents channel noise. Let $\hat{\Theta}$ denote the receiver's estimate of Θ and $P_e = P(\hat{\Theta} \neq \Theta)$ the probability of error. Construct a Bayesian estimator that minimizes P_e. Give the decision rule in as simple a form as possible. Calculate the resulting P_e as a function of $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du$, the CDF of standard normal. Evaluate the result for $A = 4$, $\sigma^2 = 4$. (Use the table in Problem 2.)

$P_e = 0.0228$

(b) (5 pts) Let Θ be uniformly distributed over the interval $[4, 10]$ and let $X = \Theta + W$ be a measurement of Θ where the measurement noise W is uniformly distributed over $[-1, 1]$. Find the Linear Least Mean Squares (LLMS) estimator of Θ and the associated mean square error MSE.

$\hat{\Theta} = \frac{9X+7}{10}$

$\text{MSE} = 0.3$

Use this space to show your work for P4 **only**.

---

a) $f_{X_1, X_2, X_3, \ldots, X_n}(x_1, x_2, x_3, \ldots, x_n) = (\theta e^{-\theta x_1})(\theta e^{-\theta x_2}) \ldots (\theta e^{-\theta x_n})$

$$= \theta^n e^{-\theta(x_1 + x_2 + \ldots + x_n)}$$

To maximize this, we take the derivative with respect to θ,

$$n\theta^{n-1} e^{-\theta(x_1+x_2+x_3+\ldots+x_n)} - (x_1+x_2+x_3+\ldots+x_n) e^{-\theta(x_1+x_2+x_3+\ldots+x_n)} \theta^n = 0$$

$$\Rightarrow n\theta^{n-1} = (x_1+x_2+x_3+\ldots+x_n)\theta^n \Rightarrow n = (x_1+x_2+x_3+\ldots+x_n)\theta$$

$$\Rightarrow \hat{\theta}_n = \frac{n}{(x_1+x_2+\ldots+x_n)} = \frac{n}{\sum_{i=1}^{n} x_i}$$

$$\lim_{n\to\infty} \hat{\theta}_n = \lim_{n\to\infty} \left(\frac{n}{\sum_{i=1}^{n} x_i}\right) = \lim_{n\to\infty} \left(\frac{\sum_{i=1}^{n} x_i}{n}\right)^{-1} = \left(\lim_{n\to\infty} \frac{\sum_{i=1}^{n} x_i}{n}\right)^{-1}$$

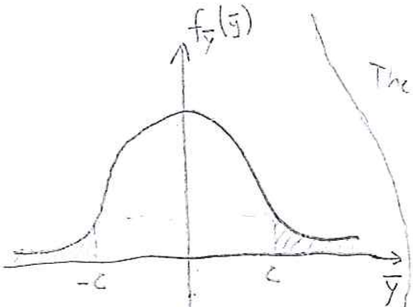
By weak law of large numbers,

$$\left(\lim_{n\to\infty} \frac{\sum_{i=1}^{n} x_i}{n}\right)^{-1} \overset{(a)}{=} (E[X])^{-1} = \frac{1}{E[X]} = \theta \Rightarrow \text{consistent}$$

$\left(\lim_{n\to\infty} P(|\hat{\theta}_n - \theta| > \epsilon) = 0 \text{ for all } \epsilon\right)$

$\boxed{\text{In step (a), } M_n = E[X] \text{ is used}}$

b) Let $\bar{X}$ be, the mean for 100 samples and $\bar{Y} = \frac{\bar{X} - 5}{1/10}$. ($\bar{Y}$ is standard normal, because $\mu = 5$, $\sigma = \frac{1}{10}$ for M_{100})



The probability of shaded area is to be 5% (2.5% each)

$$\Phi^{-1}(1 - 2.5\%) = \Phi^{-1}(0.975) = 1.96 = c$$

$\Rightarrow$ Acceptance region: $-1.96 \leq \bar{Y} \leq 1.96$

$$-1.96 \leq \frac{\bar{X} - 5}{1/10} \leq 1.96 \Rightarrow -0.196 \leq \bar{X} - 5 \leq 0.196 \Rightarrow \bar{X} \in [4.804, 5.196]$$

P4. (10 points) The two parts of this problem are independent.

(a) (5 pts) A certain type of light bulb has a lifetime X that is exponential with parameter $\theta > 0$: $f_X(x;\theta) = \theta e^{-\theta x}$, $x \geq 0$. Let $x_1, x_2, \ldots, x_n$ be the measured lifetimes of n such light bulbs. We model these measurements are samples of random variables $(X_1, \ldots, X_n)$ with a joint PDF such that

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n; \theta) = \prod_{i=1}^{n} f_X(x_i; \theta).$$

Determine the ML (maximum-likelihood) estimator $\hat{\Theta}_n(X_1, \ldots, X_n)$. Is the estimator consistent? Justify.

$\hat{\Theta}_n(X_1, \ldots, X_n) = \dfrac{n}{\sum_{i=1}^{n} X_i}$ ✓

Consistent? Yes, justification is given in the solution page.

(b) (5 pts) Let X be a normal random variable with mean μ and unit variance. We want to test the hypothesis $H_0 : \mu = 5$ at the 5% significance level, using $n = 100$ independent samples of X. What is the range of values of the sample mean $M_n = (X_1 + \cdots + X_n)/n$ for which the hypothesis is accepted? Show your work in detail. In particular show how you are using the CDF of the unit normal $\Phi(x)$.

Range of M_n = $[4.804, 5.196]$ ✓