The first 12 questions are multiple choice. Circle the correct answer. Circle only one answer.

1. (5 points) A random variable $\boldsymbol{X}$ has PMF $\Pr[\boldsymbol{X}=k]=[0.4,0.3,0.2,0.1]$ for $k=0,1,2,3$. What is the variance of $\boldsymbol{X}$?
   (a) 0
   (b) 0.5
   (c) 1
   (d) 1.5
   (e) 2

   $EX = 0\times 0.4 + 1\times 0.3 + 2\times 0.2 + 3\times 0.1 = 1.0$
   $EX^2 = 0\times 0.4 + 1^2\times 0.3 + 2^2\times 0.2 + 3^2\times 0.1 = 2.0$
   $\mathrm{Var}\, X = 2.0 - 1.0^2 = 1.0$

2. (5 points) A random variable $\boldsymbol{X}$ has PMF $\Pr[\boldsymbol{X}=k]=[0.5,0.1,0.1,0.1,0.1,0.1]$ for $k=0,1,2,3,4,5$. What is $E(\boldsymbol{X})$?
   (a) 0
   (b) 0.5
   (c) 1
   (d) 1.5
   (e) 2

   $EX = 0\times 0.5 + 1\times 0.1 + 2\times 0.1 + 3\times 0.1 + 4\times 0.1 + 5\times 0.1$
   $= 1.5$

3. (5 points) The MGF of $\boldsymbol{X}$ is $(pe^u+q)e^{2u}$ where $q=1-p$. What is $E(\boldsymbol{X})$?
   (a) p
   (b) e^u
   (c) p^2
   (d) $2+p$
   (e) $3+p$

   $EX = \frac{\partial}{\partial u} m(u)\Big|_{u=0} = \frac{\partial}{\partial u}(pe^{3u}+qe^{2u})\Big|_{u=0} = 3pe^{u}+2qe^{u}\Big|_{u=0}$
   $= 3p+2q = p+2$

4. (5 points) Two ordinary six sided dice are rolled (the dice are independent and the sides are numbered from 1 to 6). Let $\boldsymbol{S}$ the sum of the two dice. What is $\Pr[\boldsymbol{S}=5]$?
   (a) 1/3
   (b) 1/2
   (c) 1/6
   (d) 5/6
   (e) 1/9

   $P[S=5] = P[(1,4) \text{ or } (2,3) \text{ or } (3,2) \text{ or } (4,1)]$
   $= \frac{1}{36}+\frac{1}{36}+\frac{1}{36}+\frac{1}{36} = \frac{4}{36} = \frac{1}{9}$

5. (5 points) Assume $\boldsymbol{X}$ is Bernoulli with parameter p and let $\boldsymbol{Y}$ be defined by the conditional probabilities $\Pr[\boldsymbol{Y}=k|\boldsymbol{X}=k]=1-\epsilon$ and $\Pr[\boldsymbol{Y}=k|\boldsymbol{X}=1-k]=\epsilon$ for $k=0$ and $k=1$. What is $\Pr[\boldsymbol{Y}=1]$?

$P[Y=1] = P[Y=1|X=1]\,P[X=1] + P[Y=1|X=0]\,P[X=0]$
$= (1-\varepsilon)p + \varepsilon(1-p)$

   (a) $(1-\epsilon)p+\epsilon(1-p)$
   (b) $\epsilon p+(1-\epsilon)(1-p)$
   (c) ϵp
   (d) $1-\epsilon p$
   (e) $1+\epsilon p$

6. (5 points) Let $\boldsymbol{N}$ have a Poisson distribution with parameter λ. What is $\Pr[\boldsymbol{N}=k]$ for $k=0,1,2,\ldots$?
   (a) $1/k$
   (b) $\frac{\lambda^k e^{-\lambda}}{k!}$
   (c) $(1-\lambda)\lambda^k$
   (d) $\binom{N}{k}\lambda^k(1-\lambda)^{N-k}$
   (e) $\frac{e^{-\lambda k}}{1-e^{-\lambda}}$

7. (5 points) If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are independent with PMFs $[1/2, 1/4, 1/8, 1/16, \ldots]$ and $[1/2, 1/2]$, respectively, what are the first three terms of the PMF of $\boldsymbol{S}=\boldsymbol{X}+\boldsymbol{Y}$?
   (a) $[1/2, 1/4, 1/8]$
   (b) $[1/4, 3/8, 3/16]$
   (c) $[1/4, 1/4, 1/8]$
   (d) $[1/4, 1/2, 1/8]$
   (e) $[1/4, 1/8, 1/16]$

	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	
$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	
$\frac{1}{2}$		$\frac{1}{4}$	$\frac{1}{8}$	
	$\frac{1}{4}$	$\frac{3}{8}$	$3/16$	

8. (5 points) If $\boldsymbol{X}\sim U(0,1)$ what is the distribution of $\boldsymbol{Y}=2\boldsymbol{X}+3$?
   (a) $U(0,1)$
   (b) $U(0,5)$
   (c) $U(2,3)$
   (d) $U(2,5)$
   (e) $U(3,5)$

when $X=0$, $Y=2\times0+3=3$
when $X=1$, $Y=2\times1+3=5$

9. (5 points) If $\boldsymbol{X}$ has mean 1 and variance 5 and $\boldsymbol{Y} = 2\boldsymbol{X}^2 + 3$, what is $E(\boldsymbol{Y})$?
   (a) 5
   (b) 10
   (c) 15
   (d) 20
   (e) 25

$EY = E(2X^2+3) = 2E(X^2) + 3$
$= 2(1^2+5)+3 = 15$
$EX^2 = \mathrm{Var}\,X + \mu^2$

10. (5 points) If $\boldsymbol{X} \sim N(1,4)$ and $\boldsymbol{Y} = 2\boldsymbol{X} + 1$, what is $\Pr[\boldsymbol{Y} < 2]$?
   (a) 0.1587
   (b) 0.2266
   (c) 0.3085
   (d) 0.4013
   (e) 0.5000

$P(Y<2) = P(2X+1<2) = P(X<\frac{1}{2})$
$= P\left(\frac{X-1}{2} \le \frac{\frac{1}{2}-1}{2}\right) = \Phi(-\frac{1}{4})$
$= 1-\Phi(\frac{1}{4}) = 1-0.5987 = 0.4013$

11. (5 points) What is the median of the following samples: 1, 2, 1, 5, 10, 6, 4?
   (a) 1
   (b) 2
   (c) 4
   (d) 6
   (e) 10

Sort 1 1 2 4 5 6 10

12. (5 points) Many people perform 1000 independent flips of a fair coin and count the number of heads. About 2/3 of the counts will fall into which range:
   (a) (300, 700)
   (b) (400, 600)
   (c) (450, 550)
   (d) (468, 532)
   (e) (484, 516)

let N = # flips
N is binomial with params $p = \frac{1}{2}$ and $n = 1000$

$EN = np = 500$
$\mathrm{Var}\,N = npq = 250 \approx 16^2$
about $\frac{2}{3}$ will lie between $\mu-\sigma$ and $\mu+\sigma$

c d d e
a b b e
c d c e

The next six questions are short answer questions.

13. (5 points) Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be independent with $\boldsymbol{X} \sim N(\mu_x, \sigma_x^2)$ and $\boldsymbol{Y} \sim N(\mu_y, \sigma_y^2)$. What is the joint density of $\boldsymbol{X}$ and $\boldsymbol{Y}$?

$$f_{XY}(x,y) = f_X(x) f_Y(y) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2}\left[\left(\frac{X-\mu_x}{\sigma_x}\right)^2 + \left(\frac{Y-\mu_y}{\sigma_y}\right)^2\right]}$$

14. (5 points) If $\boldsymbol{X}_i$ for $i = 1, 2, \ldots, n$ are IID $U(0,1)$, what is the approximate distribution of $\boldsymbol{S} = \boldsymbol{X}_1 + \boldsymbol{X}_2 + \cdots + \boldsymbol{X}_n$?

By CLT, $S \sim N(n\mu, n\sigma^2)$

$\mu = EX = \frac{1}{2}$

$\sigma^2 = Var X = \frac{1}{12}$

$S \sim N(\frac{n}{2}, \frac{n}{12})$

15. (5 points) What is the Law of Total Probability? (write it in terms of events).

$$P(A) = P(A|B)P(B) + P(A|\bar{B})P(\bar{B})$$

16. (5 points) Let $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, …, $\boldsymbol{X}_n$ be n IID random variables with CDF $F_{\boldsymbol{X}}(x)$. Let $\boldsymbol{Y} = \max(\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n)$, i.e., $\max(1,2,4,3) = 4$. What is the CDF of $\boldsymbol{Y}$?

$$F_Y(y) = P(Y \le y) = P(\max(X_1, \ldots, X_n) \le y)$$
$$= P(X_1 \le y \text{ and } X_2 \le y \text{ and } \cdots \text{ and } X_n \le y)$$
$$= P(X_1 \le y)\, P(X_2 \le y) \cdots P(X_n \le y)$$
$$= F_X(y)^n$$

17. (5 points) Let $\boldsymbol{X}$ be exponential with parameter λ. What is $\Pr[\boldsymbol{X} > x_0]$?

$$P(X > 0) = \int_{x_0}^{\infty} \lambda e^{-\lambda x}\, dx = e^{-\lambda x_0}$$

18. (5 points) Let $\boldsymbol{N}$ be binomial with parameters $n = 5$ and $p = 1/4$. Give an expression for $\Pr[\boldsymbol{N} \le 2]$.

$$P(N \le 2) = P(N=0) + P(N=1) + P(N=2)$$
$$= \binom{5}{0}\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^5 + \binom{5}{1}\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^4 + \binom{5}{2}\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^3$$
$$= \frac{243 + 405 + 270}{1024} = \frac{918}{1024} \approx 0.9$$

19. (20 points) Consider a simple hypothesis test. Under the null hypothesis H_0, the observation is $\boldsymbol{X} \sim N(-1, \sigma^2)$. Under the alternative hypothesis, $\boldsymbol{X} \sim N(1, \sigma^2)$. Consider a test based on a threshold, i.e., whether $\boldsymbol{X} > x_T$ or $\boldsymbol{X} \leq x_T$.

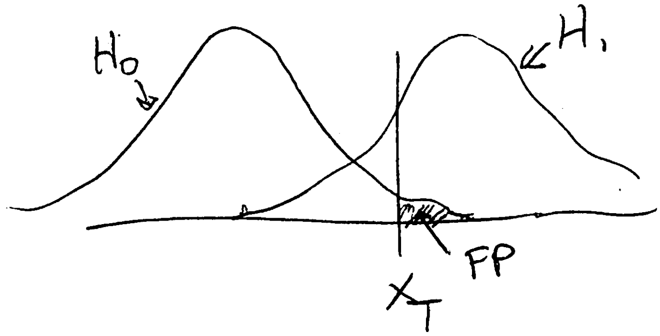
   (a) Draw two pictures, one representing the false positive probability and the second representing the false negative probability. Clearly label the two pictures.

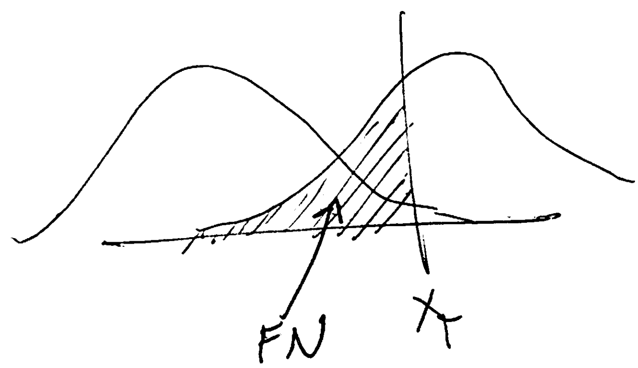
   (b) For a test comparing $\boldsymbol{X}$ to a threshold x_T, derive an expression for the false positive rate.

   (c) For a test comparing $\boldsymbol{X}$ to a threshold x_T, derive an expression for the false negative rate.

   (d) What is the likelihood ratio of the hypothesis test. Simplify it algebraically.

a).





b) $P(FP) = P(X > X_T \mid H_0) = P\left(\frac{X+1}{\sigma} > \frac{X_T+1}{\sigma} \mid H_0\right) = 1 - \Phi\left(\frac{X_T+1}{\sigma}\right)$

c) $P(FN) = P(X < X_T \mid H_1) = P\left(\frac{X-1}{\sigma} < \frac{X_T-1}{\sigma} \mid H_1\right) = \Phi\left(\frac{X_T-1}{\sigma}\right)$

d) $$L(x) = \frac{f_1(x)}{f_0(x)} = \frac{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x+1}{\sigma}\right)^2}}{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-1}{\sigma}\right)^2}} = e^{-\frac{1}{2}\left(\frac{x+1}{\sigma}\right)^2 + \frac{1}{2}\left(\frac{x-1}{\sigma}\right)^2}$$

$$l(x) = \log L(x) = +\frac{1}{2\sigma^2}\left(x^2 - 2x + 1 - x^2 - 2x - 1\right) = \frac{2x}{\sigma^2}$$

Test reduces to $X > X_T$. for some X_T.

20. (20 points) Let $\boldsymbol{X} \sim N(1,1)$ and $\boldsymbol{Y} \sim N(-2,3)$ with $\boldsymbol{X}$ and $\boldsymbol{Y}$ independent. What are the following:

(a) $\Pr[0 \leq \boldsymbol{X} \leq 3]$

(b) $\Pr[0 \leq \boldsymbol{X}+\boldsymbol{Y} \leq 3]$

(c) $\Pr[0 \leq \boldsymbol{X}+\boldsymbol{Y} \leq 3 | \boldsymbol{Y}=2]$

(d) $\mathrm{Var}[\boldsymbol{X}-\boldsymbol{Y}]$

(e) $\mathrm{Cov}[\boldsymbol{X}-\boldsymbol{Y}, \boldsymbol{X}+\boldsymbol{Y}]$

a) $P(0 \leq X \leq 3) = P\left(\frac{0-1}{1} \leq \frac{X-1}{1} \leq \frac{3-1}{1}\right) = \Phi(2) - \Phi(-1)$

$= \Phi(2) + \Phi(1) - 1 = 0.9772 + 0.8413 - 1 \boxed{= 0.8185}$

b) $W = X+Y \quad W \sim N(1-2, 1+3) = N(-1, 4)$

$P(0 \leq W \leq 3) = P\left(\frac{0+1}{2} \leq \frac{W+1}{2} \leq \frac{3+1}{2}\right) = \Phi(2) - \Phi(\tfrac{1}{2})$

$= 0.9772 - 0.6915 = \boxed{0.2857}$

c) $P(0 \leq X+Y \leq 3 | Y=2) = P(0 \leq X+2 \leq 3) = P(-2 \leq X \leq 1)$

$= P\left(\frac{-2-1}{1} \leq \frac{X-1}{1} \leq \frac{1-1}{1}\right) = \Phi(0) - \Phi(-3)$

$= 0.5 + 0.9987 - 1$

$\boxed{= 0.4987}$

d) $Var\{X-Y\} = Var\{X\} + (-1)^2 Var\{Y\} = 1+3 = \boxed{4}$

e) $Cov[X-Y, X+Y] = Cov[X, X] + Cov(X, Y) - Cov[X, Y] - Cov[Y, Y]$

$= Var[X] + 0 + 0 - Var[Y] = 1 - 3 = \boxed{-2}$

21. (20 points) Let $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, …, $\boldsymbol{X}_n$ be a sequence of n IID observations with mean μ and variance σ^2.

(a) What is an unbiased estimate of μ?

(b) Show the mean estimate above is unbiased.

(c) What is an unbiased estimate of σ^2.

(d) Show the variance estimate above is unbiased.

a) $\overline{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$ is unbiased

b) $E\overline{X}_n = \frac{E(X_1 + \cdots + X_n)}{n} = \frac{EX_1 + EX_2 + \cdots EX_n}{n} = \frac{n\mu}{n} = \mu$

c) $\hat{\sigma^2} = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2$ is unbiased

d) $E\hat{\sigma^2} = \frac{1}{n-1} E\left(\sum_{i=1}^{n} \left(X_i^2 - 2X_i \overline{X}_n + \overline{X}_n^2\right)\right)$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n} E X_i^2 - \frac{2}{n}\sum_{i=1}^{n}\sum_{k=1}^{n} E X_i X_k + n E(\overline{X}_n^2)\right]$$

$$= \frac{1}{n-1}\left[n(\sigma^2 + \mu^2) - \frac{2}{n}(n\sigma^2 + n^2\mu^2) + \frac{n}{n^2}(n\sigma^2 + n^2\mu^2)\right]$$

$$= \frac{n-2+1}{n-1}\sigma^2 + \frac{n-2n+n}{n-1}\mu^2$$

$$= \sigma^2$$