401 Final

DO NOT START UNTIL TOLD TO DO SO.

TIME ALLOCATION WILL BE CRUCIAL AS YOU WILL LIKELY NOT HAVE ENOUGH TIME TO FINISH. FIND EASY PROBLEMS, MAKE SURE YOU SOLVE THEM.

FOR EXPLANATION PROBLEMS, CLARITY AND CONCISENESS IS SOUGHT. WRITING IRRELEVANT THINGS WILL RESULT IN POINT DEDUCTION.

DO NOT PANIC.

DO NOT WRITE YOUR RESPONSE ON THE PROBLEM SHEET. ONLY WRITE ON THE BLANK SCRATCH PAPER.

WRITE YOUR NAME ON THE FIRST PAGE OF THE PROBLEM SHEET.

WRITE YOUR NAME ON EACH SIDE OF THE BLANK SCRATCH PAPER THAT YOU WROTE A RESPONSE ON.

WRITE THE PAGE NUMBER ON EACH SIDE OF THE BLANK SCRATCH PAPER THAT YOU WROTE A RESPONSE ON IN THE FOLLOWING FORMAT. SIDE NUMBER OVER TOTAL NUMBER OF SIDES WITH YOUR WRITTEN RESPONSE.

WHEN YOU ARE DONE, SUBMIT WITH THE PROBLEM SHEET ON TOP OF YOUR WRITTEN RESPONSE.

NO NEED TO DERIVE FACTS FROM SCRATCH UNLESS EXPLICITLY DEMANDED.

GOOD LUCK.

1. (20 pts) Answer T (true) or F (false) to the following statements. No need to show intermediate derivations.

1. If you flip an unfair coin 10 times, probability of getting 3 heads is the same as that for getting 7 heads. F. Although $\binom{n}{k} = \binom{n}{n-k}$, in general $p^k(1-p)^{n-k} \neq p^{n-k}(1-p)^k$.

2. The range of values that a random variable of a geometric distribution takes is finite. F. For any large M, $P(X > M)$ is not 0.

3. A random variable with mean 1 and vairance 10 can be modelled by a Poisson distribution. The Poisson distribution has variance and mean to be equal.

4. You can approximate a Binomial distribution with a Poisson distribution if p is small and n is large. T. This is the Poisson limit theorem.

5. For the exponential distribution, the larger the mean, the larger the variance. T. Variance is the square of the mean. For positive regions, the derivative is positive.

6. If $X \sim N(5, 2)$, $P(X < 0) > 0.5$. F. Draw a diagram.

7. There exists instances of Normal distributions where getting a negative realization has 0 probability. F. The pdf of a normal is non-zero on the whole real line.

8. It is known that $P(|X| > c) = 0.1$ where X is a Normal distribution. Then $P(X < -c) = 0.05$. T. Draw a diagram. Use symmetry.

9. If $P(X > c) = d$ for $X \sim N(e, 2)$, then $P(Y > c) < d$ for $Y \sim N(e + 1, 1)$. F. Draw a diagram.

10. If you look at the pdf of a Normal, the line quickly approaches 0 as x increases or decreases. But it should never touch 0. T. Again, the pdf is never 0.

11. Joint random variable (X, Y) is supported on the unit square $[0, 1] \times [0, 1]$. It is possible that $f(x, y) > 1$ on all of $[0, 1] \times [0, 1]$. F. If it was, the integration will give a value larger than 1.

12. If a joint density can be factored, $f(x, y) = g(x)h(y)$, the marginal densities are $g(x)$ and $h(y)$. F. The constant scaling on each marginal could be off.

13. In the pmf table for bivariate random variables, the sum of values for each row should always be 1. F. The some of all values in the table should be 1.

14. The regression function, $E[Y|X = x]$ need not be linear. T. It can be nonlinear. For example try constructing an example using the exponential distribution.

15. $f_{Y|X=x}(y)$ has the same shape as $f(y, x)$ with x fixed and seen as a function of y. When we say same shape, we mean there is a constant by which we can scale one of the functions to obtain the other function. T. You only scale with the probability of $P(X = x)$.

16. The minimum as well as the maximum of two independent uniform random variables, is again an uniform random variable. F. It will no longer be an uniform. This can be seen from the fact that it still takes values from 0 to 1 but values closer to the ends will be more likely.

17. The variance of the sum of random variables X_1 up to X_n is the sum of the variances $V(X_1)$ up to $V(X_n)$. F. You also need to consider the covariance.

18. We have $f(x, y) = f(y, x)$ for the joint density. Then if $E[X] = c$, $E[Y]$ is also c. T. If you switch symbols in the integration, you get the exact same thing.

19. $Y = a + bX + \epsilon$. a and b are constants but the others are random variables. Then the variance of Y is always larger than the variance of X. F. It could be that the bX and ϵ is negatively correlated. b might be small as well.

20. The range of $f(X)$ cannot be larger than the range of X for any f. Hint, think of Uniform distributions. F. You can choose a and b suitably to have $aU + b$ have larger range.

2. (20 pts) For each of the following problems, choose the most suitable option. No need to show intermediate derivations.

1. Z is a standard normal. X is 1 if $Z > 0$ and -1 if $Z < 0$. Then,

   (a) $1 \geq Cor(X, Z) > 0$
   (b) $Cor(X, Z) = 0$
   (c) $0 > Cor(X, Z) \geq -1$
   (d) $|Cor(X, Z)| > 1$

   (a). X is high when Z is high.

2. We have two independent samples $\{X_i\}_{i=1}^{n_1}$ and $\{Y_i\}_{i=1}^{n_2}$ of X and Y. Both of them i.i.d. Which of the following is not an unbiased estimator of $\mu_X - \mu_Y$? ϵ is an irrelevant mean zero random variable.

   (a) $X_1 - Y_{n_2}$
   (b) $\bar{X} - Y_1$
   (c) $(\bar{X}, -\bar{Y})$
   (d) $\frac{\sum_{i=1}^{\min(n_1,n_2)}(X_i - Y_i + \epsilon)}{\min(n_1,n_2)}$

   (c). This is not even univariate.

3. If there are two unbiased estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ of θ, and $Cov(\hat{\theta}_1, \hat{\theta}_2) < 0$ as well as $V(\hat{\theta}_1) = V(\hat{\theta}_2)$, what is not true of $0.5\hat{\theta}_1 + 0.5\hat{\theta}_2$?

   (a) it is unbiased

(b) for linear f, $f(0.5\hat{\theta}_1 + 0.5\hat{\theta}_2)$ is unbiased for $f(\theta)$

(c) it will have larger standard error than $\hat{\theta}_1$ as well as $\hat{\theta}_2$

(d) $E[(\theta_1 - \theta)(\theta_2 - \theta)] < 0$ (c). By combining two estimators that are negatively correlated, you get smaller standard error.

4. We have a pmf table where the rows correspond to X (first row for smallest possible value of X, second row for second smallest possible value for X, etc, in order). The columns correspond to Y (first column for smallest possible value of Y, second column for second smallest possible value for Y, etc, in order). To find $F(x_5, y_7)$ where x_5 is the fifth smallest possible value of X and y_7 is the seventh smallest possible value of Y, we should add up values of the pmf table that are,

   (a) in both the first 4 rows and first 6 columns

   (b) in both the first 5 rows and first 7 columns

   (c) in the fifth row and seventh column

   (d) in both the last 5 rows and the last 7 columns

   (b). The inequality includes equality.

5. Which of the following is correct for a bivariate random variable?

   (a) $F(x, y) \geq F_X(x)$

   (b) $F_X(x) = F(x, \infty)$

   (c) $F(-\infty, y) = 0$

   (d) $F(x, y) = P(X = x, Y = y)$

   (b). It amounts to integrating out y.

6. If Y given X follows a Binomial (n, X) (X is Uniform $(0.1, 0.9)$), then,

   (a) none of the below is true

   (b) unconditional (i.e. marginal) X is not an Uniform

   (c) $E[Y|X = x]$ is nonlinear in x

   (d) $V[Y|X = x]$ is nonlinear in x

   (d). $V[Y|X = x] = nx(1 - x)$.

7. $X \sim N(\mu, \sigma^2)$. $Y = a + bX + \epsilon$ and X independent of ϵ. Which of the following is a random variable?

   (a) $E[\bar{X}]$

   (b) μ

   (c) $E[Y|X]$

   (d) $E[Y|X = x]$

(c). It's an expectation of Y as a function of X where X is random.

8. If a joint distribution (X, Y) has positive density on and only on a rectangular region,

   (a) expected value of X^Y is finite

   (b) X is independent of Y

   (c) the support of the conditional distribution $f_{X|Y=y}(x)$ (i.e. the region in the real line on which $f_{X|Y=y}(x) > 0$) is the same for any y.

   (d) two of the above is true

   (d). Conditional distributions are just looking at lines cutting through the rectangle. So (d) is true. (a) is also true. Even when $X = 0$ and $Y < 0$, this is not a big issue because it is only a line in a plane.

9. A random vector (X, Y) is uniformly distributed on the triangle with corners $(0, 0)$, $(0, 1)$, and $(1, 0)$. Which of the following is not an expression for $E[\log_X Y]$?

   (a) $\int_0^1 \int_0^{1-y} \frac{\log_x y}{2} dx dy$

   (b) $\int_0^1 \int_0^{1-x} \frac{\log_x y}{2} dy dx$

   (c) $\int_0^1 \int_0^{1-v} \frac{\log_u v}{2} du dv$

   (d) all of the above is valid expression for $E[\log_X Y]$

   (d). Draw the regions and switch symbols if necessary.

10. We have $Y = a + bX$. The correlation between Y and X is,

   (a) b

   (b) 1

   (c) b^2

   (d) $\sqrt{b}$

   (b). Follow the definitions.

3. (5 pts) Show that the discrete random variable with $P(X = k) = p(1-p)^{k-1}$ for $0 < p < 1$, $k = 1,\ 2,\ \cdots$ (the geometric distribution) has mean $\frac{1}{p}$. You can interchange infinite sum with differentiation. This derivation was done in the lecture. Please find it there. You could also use the first equation of the "expected value of the max".

4. (3 pts) You observe from the distribution in problem 3, an i.i.d. sample $X_1,\ X_2,\ \cdots,\ X_n$. Find the method of moments estimator of p. This is the reciprocal of the sample mean.

5. (4 pts) For the sample in Problem 4, show that the log likelihood function of p is $n \log p + (\sum_{i=1}^n X_i - n) \log(1 - p)$. Just take the log of $\Pi_{i=1}^n p^n (1 - p)^{X_i - n}$.

6. (4 pts) Using Problem 4, show that the maximum likelihood estimator of p is the same as the method of moments estimator. Take the derivative of the log-likelihood above, set it to 0 and solve for p.

7. (3 pts) We will use $\bar{X}$ as the estimator of λ for an i.i.d. sample from the exponential distribution with density $\frac{1}{\lambda}\exp(-\frac{x}{\lambda})$ for $x \geq 0$. Show that it is unbiased. Use that the mean of X is λ. You just need to show that the sample mean also has mean λ. This is also in the lecture.

8. (3 pts) Calculate the standard error of the estimator in problem 7. i.e. calculate the standard deviation of $\bar{X}$. Use that $V[X] = \lambda^2$ and the fact that the samples are i.i.d.

9. (4 pts) Find the mean squared error of the estimator in problem 7. i.e. calculate $E_\lambda(\bar{X} - \lambda)^2$. Use that the MSE can be decomposed into the variance and the bias squared.

10. (3 pts) X is Poisson with density $\frac{e^{-\lambda_1}\lambda_1^k}{k!}$ for $k = 0,\ 1,\ \cdots$. Y is Poisson with density $\frac{e^{-\lambda_2}\lambda_2^k}{k!}$ for $k = 0,\ 1,\ \cdots$. X and Y are independent. Assume $\lambda_1,\ \lambda_2 > 0$. We have

$$P(X+Y=k) = \sum_{i=0}^{k} P(Y=k-i)P(X=i) \tag{1}$$

$$= \sum_{i=0}^{k} P(Y=i)P(X=k-i) \tag{2}$$

Explain in words what this equality means. The probability that the sum equals k is the sum of the probabilities that $X = i$ where i from 0 to k and Y is equal to $k - i$.

11. (5 pts) In the setting of problem 10, use equation (1) to show that

$$P(X+Y=k) = \frac{e^{-\lambda_1-\lambda_2}(\lambda_1+\lambda_2)^k}{k!}.$$

Note that $\frac{\lambda_1}{\lambda_1+\lambda_2} + \frac{\lambda_2}{\lambda_1+\lambda_2} = 1$ and $0 < \frac{\lambda_1}{\lambda_1+\lambda_2},\ \frac{\lambda_2}{\lambda_1+\lambda_2} < 1$. You must use that the binomial probabilities add up to 1.

12. (3 pts) What can you conclude about the random variable $X + Y$, from the results of problem 11? It follows a poisson.

13. (5 pts) Show that

$$P(X \leq s+t | X \geq s) = 1 - \exp(-\lambda t)$$

where X is exponential with density $\lambda\exp(-\lambda x)$ for $x \geq 0$ and $\infty > \lambda > 0$. This was a homework problem.

14. (3 pts) From the result of problem 13, what can you conclude about the distribution of $X - s$ given $X \geq s$? It follows an exponential.

15. (3 pts) $X_1, \cdots, X_n$ is an i.i.d. sample form a random variable with density

$$\frac{1}{2\sqrt{\pi}} \exp\left(-\frac{1}{4}[x-3]^2\right).$$

What is the mean and variance of this random variable? Mean is 3, variance is $\sqrt{2}$.

16. (3 pts) What distribution does $\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}$ follow? Where μ is the mean and σ is the square root of the variance you found in problem 15. Standard Normal.

17. (3 pts) What distribution does $\frac{\sqrt{n}(\bar{X}-\mu)}{S}$ follow? Where

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

t with $n-1$ degrees of freedom.

18. (6 pts) Show that the expected value of S^2 is the variance you found in problem 15. This is in the lecture notes, refer there.