# E1: Midterm Exam

## CEE 260/MIE 273: Probability & Statistics in Civil Engineering

October 8, 2020

time limit: Twenty-Four Hours

---

## Name

Please print your name clearly in the box below.

| |
|---|
| |

*Turn to the next page to read the instructions.*

## Instructions

This exam contains **15 pages** (including the front and back pages) and **8 problems, 54 points** (a final question is entirely extra credit; 4pts). You have **24 hours** to complete it. You may print out the PDF, complete it and upload as a PDF on Moodle, or *neatly* answer the questions on blank pages of paper, scan and upload.

This is an **open resource examination**. You are expected to complete the exam individually. Asking anyone (colleague, friend, tutor, etc) questions about the exam is *not allowed*. If any questions arise during the exam, direct them to me (via email).

The following rules apply:

- **Organize your work**, in a reasonably neat and coherent way.Work scattered all over the page without a clear ordering will receive very little credit.
- **Show ALL your work where appropriate**. The work you show will be evaluated as well as your final answer. Thus, provide ample justification for each step you take. Indicate when you have used a probability table or MATLAB to obtain a result. In the case of MATLAB, briefly include the function or statement you used to arrive at your result. In the long response questions, simply putting down an answer without showing your steps will not merit full credit. **EXCEPTION:** For short response or "True/False" questions, *no explanations are required*. However, the more work you show, the greater your chance of receiving partial credit if your final answer is incorrect.
- If you need more space, use the blank pages at the end, and clearly indicate when and where you have done this.
- Questions are roughly in order of the lectures, so later questions may not necessarily be harder. If you are stuck on a problem, it may be better to skip it and get to it later.
- Manage your time wisely.

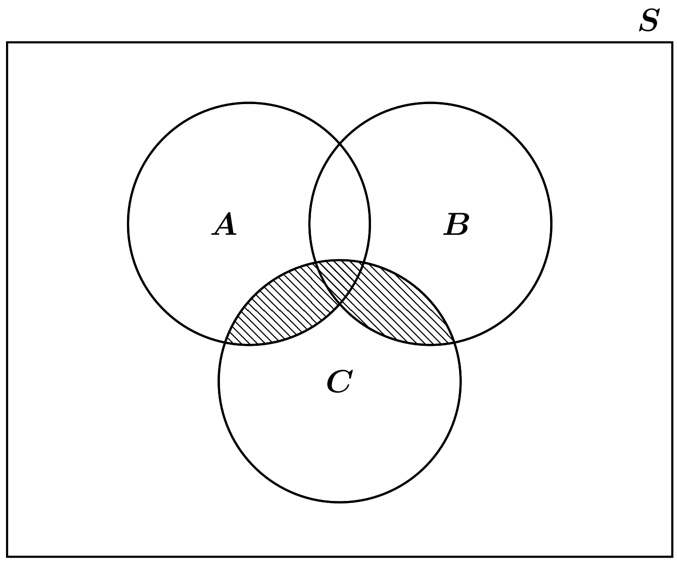
## Problem 1 *True/False questions (11 points)*

Respond "T" (*True*) or "F" (*False*) to the following statements.

| | | |
|---|---|---|
| **(i)** | T | The number of heads in 50 tosses of a coin is a discrete random variable. |
| **(ii)** | F | The number of car accidents at a busy intersection within a given time interval is a continuous random variable. |
| **(iii)** | F | The time interval between successive car accidents at a busy intersection is a discrete random variable. |
| **(iv)** | T | $P(AB) = 1 - P(\overline{AB})$ for any two events $A$ and $B$. |
| **(v)** | F | If $A$ and $B$ are independent, then $P(B\|A) = P(A)$. If independent, then $P(B\|A) = P(B)$ and $P(A\|B) = P(A)$. |
| **(vi)** | T | If three events $A$, $B$ and $C$ are collectively exhaustive, then $P(A \cup B \cup C) = 1$. |
| **(vii)** | T | The maximum value of a CDF curve is equal to 1. |
| **(viii)** | T | A binomial distributed variable $X$ with sufficiently large $n$ can be approximated by a Poisson distribution whose variance is given by $Var(X) = np$. |
| **(ix)** | T | The standard normal variate $Z$ has a variance of 1. |
| **(x)** | T | The mean of a lognormal distribution is always greater than its median. |
| **(xi)** | T | If a variable $X$ is normally distributed, then $e^X$ is lognormally distributed. |

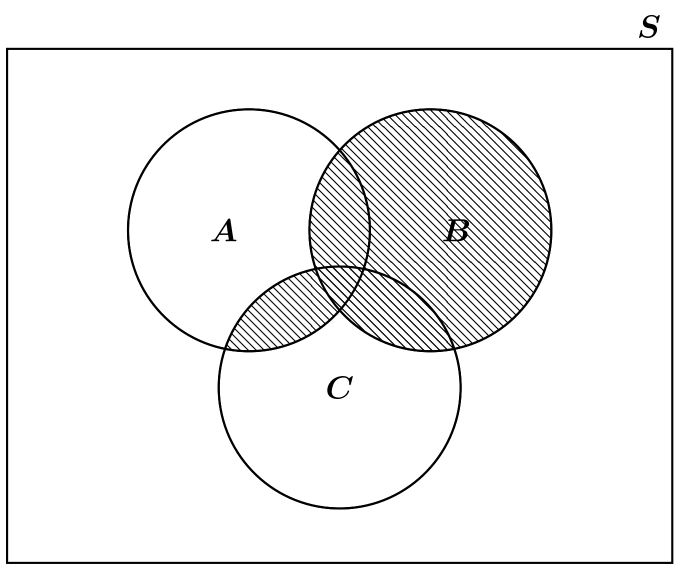
## Problem 2 *Short answer questions (19 points)*

*[2]* **(i)** Write the combination of events (using set notation) depicted in the figure below.



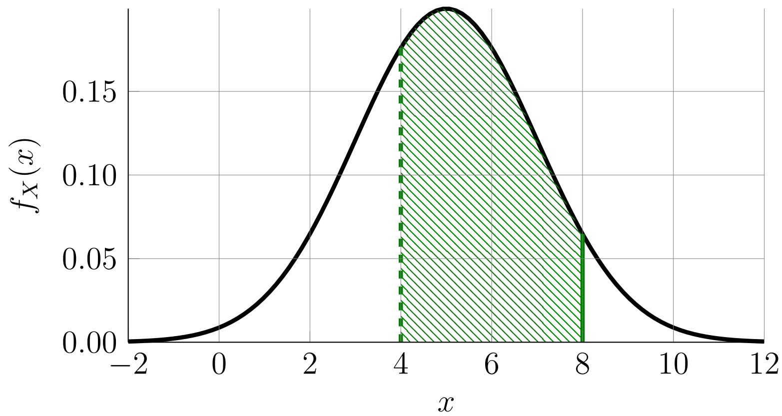

**Answer:** $AC \cup BC \equiv (A \cup B) \cap C$

*[2]* **(ii)** Write the combination of events (using set notation) depicted in the figure below.

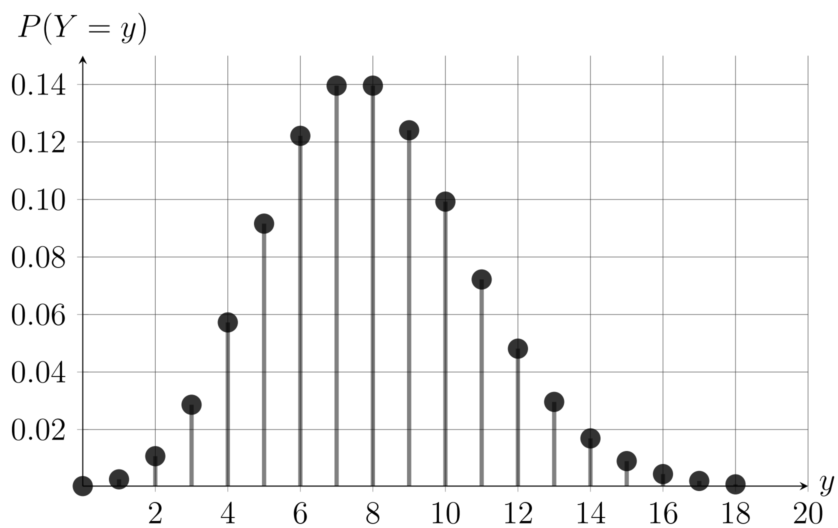



**Answer:** $B \cup AC$

**(iii)** Shade the area under the curve that gives you the probability $P(X > 4 \cap X \leq 8)$. *[2]*

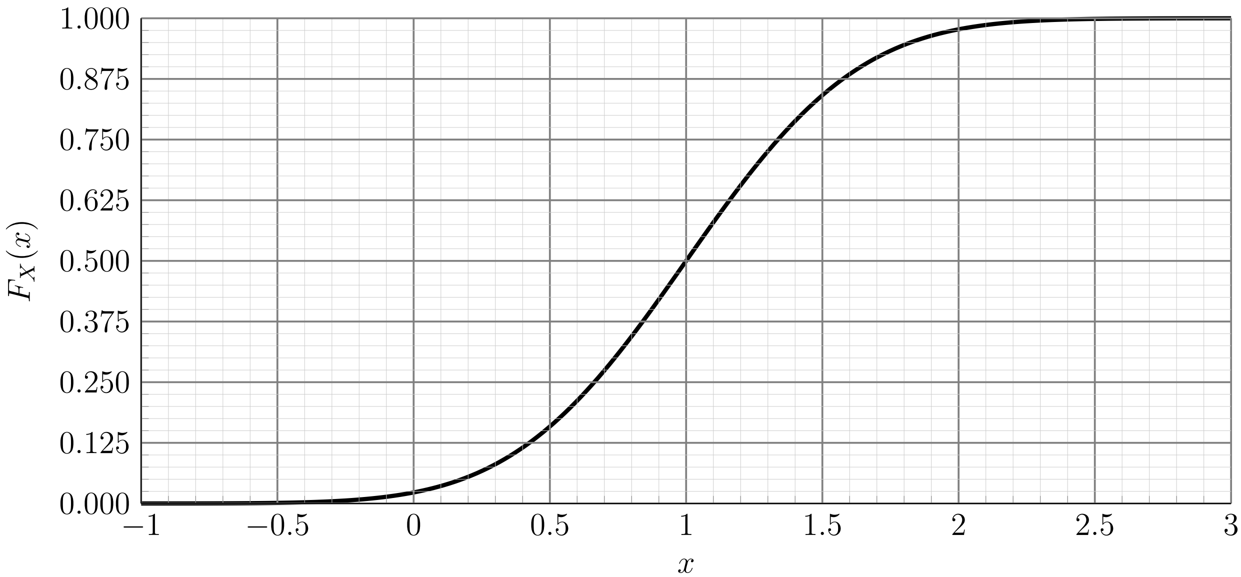



**(iv)** The PMF of a random variable $Y$ is given in the figure below. Use the figure to estimate the probability $P(Y = 8 \cup Y = 10)$. *[2]*




**Answer:** $P(Y = 8 \cup Y = 10) \approx 0.14 + 0.10 = 0.24$

[1] **(v)** Below is the CDF of a given normal distribution. Estimate the second quartile of the distribution.




**Answer:** $Q2 \approx 1$ (median)

**(vi)** Estimate the interquartile range (IQR) of the distribution plotted in the figure above.

[2]

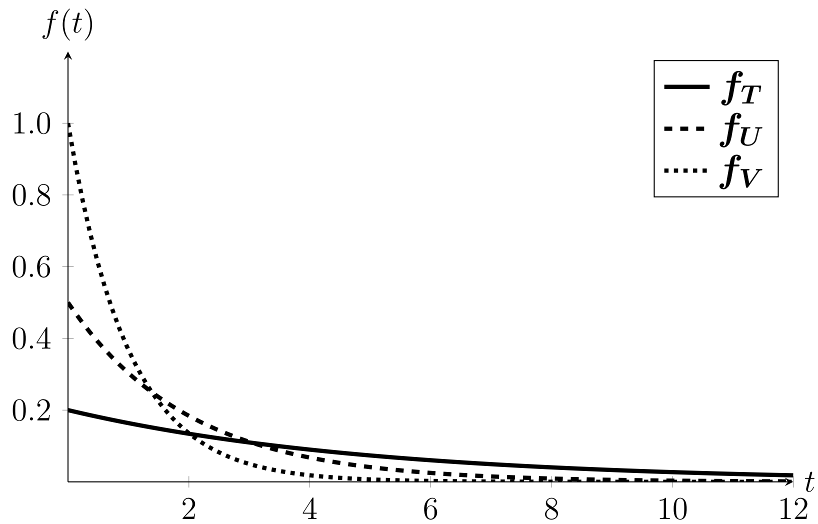
**Answer:** $IQR = Q3 - Q1 \approx 1.35 - 0.65 = 0.70$

**(vii)** The median of a lognormal variate is 6. What is the mean of the associated normal variate?

[2]

**Answer:** This is simply the parameter $\mu = \ln(6) = 1.79$

**(viii)** Consider the PDFs of the epxonential random variates $T$, $U$ and $V$ (measured in hours) shown in the figure below. Which of them has the lowest mean? *[1]*



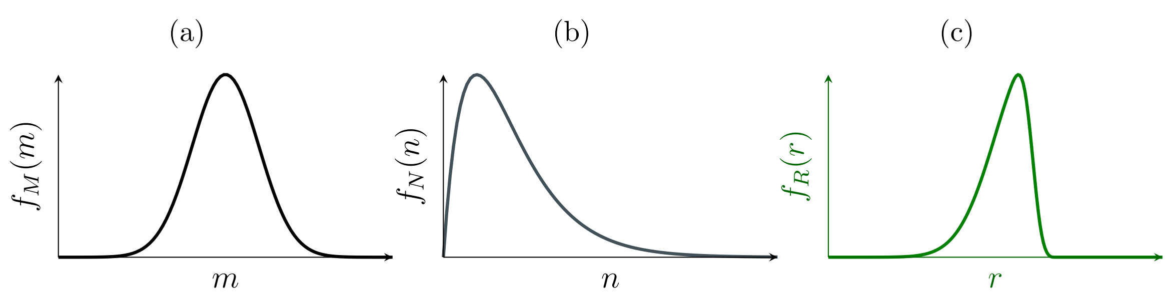

**Answer:** The mean of an exponential r.v. is the inverse of $\lambda$. Thus, $V$ has the smallest mean since $\lambda_V = 1$hr.

**(ix)** In the figure above, which random variable has mean elapsed time (or mean interval time) of 2 hours? *[1]*

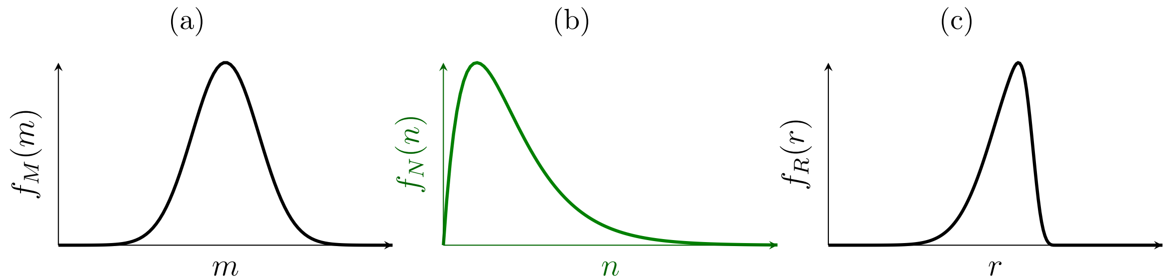
**Answer:** $\lambda_U = 1/0.5 = 2$. Thus $U$ has a mean of 2 hrs.

**(x)** Which of the distributions below is left-skewed? Circle (a), (b) or (c). *[1]*

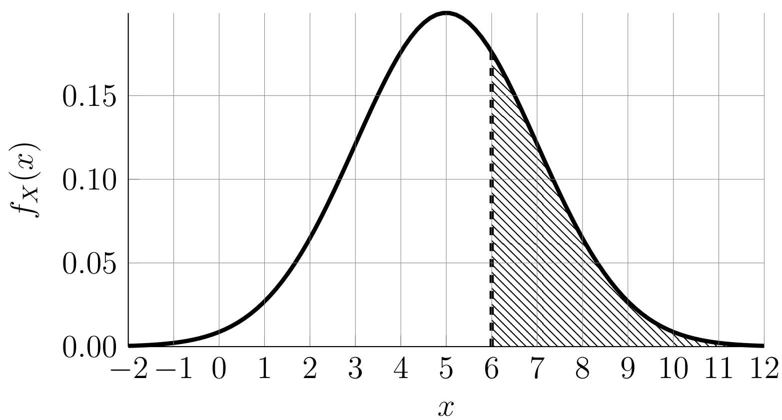
This is (c), as it has a long left tail.



*[1]* **(xi)** In which of the distributions is the mean greater than the median? Circle (a), (b) or (c). The correct answer is (b) as it is a right-skewed distribution.




*[2]* **(xii)** $X \sim N(\mu = 5, \sigma^2 = 4)$. Find the probability indicated by the shaded portion of the PDF below.




**Answer:** $P(X > 6) = 1 - P(X \le 6) = 1 - \Phi\left(\frac{6-5}{\sqrt{4}}\right) = 1 - \Phi(0.5) = 0.309$

## Problem 3 *Events and probability (3 points)*

You are given the following probabilities of three events: *[3]*

$$P(A) = 0.1$$

$$P(B) = 0.1$$

$$P(B|A) = 0.4$$

Find $P(A|B)$.

First we write

$$P(A|B) = \frac{P(AB)}{P(B)}$$

Next we find $P(AB)$:

$$P(AB) = P(BA) = P(B|A)P(A) = 0.4(0.1) = 0.04$$

Therefore,

$$P(A|B) = \frac{0.04}{0.1} = \boxed{0.4}$$

Alternatively, you can observe that $A$ and $B$ have equal probability. In which case, $P(A|B) = P(B|A)$.

## Problem 4 *Binomial distribution (8 points)*

85% of all vehicles examined at an emissions inspection station pass. Successive vehicles pass or fail independently of one another.

**(a)** Find the probability that all of the next five vehicles *pass* inspection. *[2]*

Let $p = 0.85$ (probability of passing). Based on the fundamental principle of counting, the probability that all next 5 cars pass is:

$$P(X = 5) = p^5 = 0.85^5 = \boxed{0.4437}$$

*[2]* **(b)** If only three vehicles in a batch of six vehicles pass inspection, how many ways can this occur?

The number of ways in which 3 vehicles can pass out of 6 of them is given by:

$$\binom{6}{3} = \frac{6!}{3!(6-3)!} = \frac{6 \times 5 \times 4 \times 3!}{3! \times 3!}$$
$$= \boxed{20}$$

*[1]* **(c)** If a batch of 7 vehicles are inspected, how many do you expect to *fail*?

The expected number of failures in 7 vehicles, where the probability of failure is $1 - p = 0.15$ is given by:

$$n(1-p) = 7(0.15) = \boxed{1.05}$$

Approximately 1 in 7 vehicles will fail on average.

*[3]* **(d)** What is the probability that at most 1 out of 7 vehicles will *fail* inspection?

There are several ways to formulate the solution to this question. However, we can simply use the binomial formula and replace $p$ with 0.15 (since the event of interest is failing inspection, not passing).

$$\begin{aligned} P(X \le 1) &= \sum_{k=0}^{1} \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=0}^{1} \binom{7}{k} 0.15^k (0.85)^{n-k} \\ &= \binom{7}{0} (0.15)^0 (0.85)^7 + \binom{7}{1} (0.15)^1 (0.85)^6 \\ &= 0.85^7 + 7(0.15)(0.85)^6 = \boxed{0.7166} \end{aligned}$$

In MATLAB, you can also use `binocdf(1, 7, .15)` to obtain your answer.

# Problem 5 *Poisson distribution (3 points)*

Occasionally an airline will lose a bag. Suppose a small airline has found it can reasonably model the number of bags lost each weekday using a Poisson model with a variance of 2.2 bags$^2$. What is the probability that the airline will lose 1 or 2 bags next Monday? *[3]*

Let $X$ be the number of bags lost per day.

Thus, $X \sim Poisson(\lambda)$. For a Poisson distribution, the mean and the variance are both equal to $\lambda$. Thus,

$$\lambda = 2.2$$

The PMF of the Poisson distribution is given by:

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

Therefore, the required probability is given by

$$\begin{aligned} P(X = 1 \cup X = 2) &= \frac{2.2^1 e^{-2.2}}{1!} + \frac{2.2^2 e^{-2.2}}{2!} \\ &= 2.2e^{-2.2}(1 + 1.1) \\ &= 2.1(2.2)(e^{-2.2}) = \boxed{0.512} \end{aligned}$$

In MATLAB, you can obtain this result using `poisspdf(1, 2.2) + poisspdf(2, 2.2)` or `poisscdf(2, 2.2) - poisscdf(0, 2.2)`.

## Problem 6 *Normal distribution (3 points)*

*[4]* The average daily high temperature in June in LA is 77°F with a standard deviation of 5°F. Suppose that the temperatures in June closely follow a normal. How hot are the hottest 15% of the days during June in LA?

We want to find the 85th percentile of the data here. In other words, what is the temperature $x$, where $P(X > x) = 0.85$?
Thus, we find the appropriate $Z$-score and convert to $x$.

$$z = \Phi^{-1}(0.85) = 1.0364$$

Given that $z = \frac{x-\mu}{\sigma}$, then

$$x = z\sigma + \mu = 1.0364(5) + 77 = 82.2$$

The hottest 15% of the days in June in LA are at least 82.2°F.
(Note that you can also arrive at this answer in MATLAB using `norminv(.85,77,5)`.)

## Problem 7 *Logormal distribution (3 points)*

Given that the lifetime in days of an electronic component is lognormally distributed with median 3 days and $\sigma = 0.3$. Find the probability that a component lasts less than 2 days.

The median $x_m$ of a lognormal distribution is $e^{\mu}$. Thus,

$$\mu = \ln(x_m) = \ln(3)$$

The required probability is therefore

$$P(X < 2) = \Phi\left(\frac{\ln(2) - \ln(3)}{0.3}\right) = \Phi(-1.3516) = \boxed{0.0883}$$

You can also solve for this in MATLAB using `logncdf(2,log(3),.3)`.

## Problem 8 *Exponential distribution (4 points)*

The delay time $T$ of a flight is exponentially distributed with $\lambda = 4$ (mean rate of occurrence per hour).

(a) What is the expectation of $T$? *[1]*

This is the average "waiting time".

$$E(T) = \frac{1}{\lambda} = 0.25 \text{ hr}$$

(b) What is the standard deviation of $T$? *[1]*

The standard deviation of $T$ is

$$SD(T) = \frac{1}{\lambda} = 0.25 \text{ hr}$$

(c) What is the probability that a flight is delayed by at least 30 minutes? *[2]*

$\lambda$ is specified with respect to the hour. So, we convert 30 minutes to 0.5 hr.

$$\begin{aligned} P(T \geq 0.5) &= 1 - (1 - e^{-\lambda t}) = e^{-\lambda t} \\ &= e^{-4(0.5)} = e^{-2} \\ &= \boxed{0.135} \end{aligned}$$

You can solve for this in MATLAB via `expcdf(0.5, 0.25, 'upper')`.

## Problem 9 *Joint distributions (Extra Credit, 4 points)*

The joint PMF of two random variables $X$ and $Y$ is shown in the table below.

| | | $y$ | | |
|---|---|---|---|---|
| $p_{X,Y}(x,y)$ | | **0** | **1** | **2** |
| | **0** | 0.10 | 0.04 | 0.02 |
| $x$ | **1** | 0.08 | 0.20 | 0.06 |
| | **2** | 0.06 | 0.14 | 0.30 |

*[2]* **(i)** Find $P(\overline{X = 1 \cap Y = 2})$.

Find the cell where $X = 1$ and $Y = 2$. Subtracting the value here from 1 gives the complement.

$$P(X = 1 \cap Y = 2) = 0.06$$
$$P(\overline{X = 1 \cap Y = 2}) = 1 - 0.06 = \boxed{0.94}$$

*[2]* **(ii)** Compute $P(X = 0|Y < 1)$.

The event $Y < 1$ is the first column of the table. The event $X = 0$ is the first cell in that column.

Thus,

$$\begin{aligned} P(X = 0|Y < 1) &= \frac{P(X = 0 \cap Y < 1)}{P(Y < 1)} \\ &= \frac{0.10}{0.10 + 0.08 + 0.06} \\ &= \frac{0.10}{0.24} \\ &= \boxed{0.417} \end{aligned}$$

CEE 260/MIE 273 | J. B. Oke | Fall 2020

Department of Civil and Environmental Engineering

University of Massachusetts Amherst