On my honor, I have neither received nor given any unauthorized assistance on this examination. Signed by Ethan Wong.
Ethan Wong

Ethan Wong

6/7/21

CEE 110 Final

1. Uniformly distributed between 95 and 103 minutes

$$f(x;A,B) = \begin{cases} \frac{1}{103-95} & 95 \le x \le 103 \\ 0 & \text{else} \end{cases}$$

a) $\mu_x = \int_A^B x \frac{1}{8} dx$

$\mu_x = \int_{95}^{103} \frac{1}{8} x \, dx = \frac{1}{8}\left[\frac{x^{1+1}}{2}\right]_{95}^{103} = 99$

$\boxed{\text{The mean is 99}}$

b) $V(x) = E(x^2) - [E(x)]^2$

$E(x^2) = \int_{95}^{103} x^2 \frac{1}{8} dx = \frac{1}{8}\left[\frac{x^3}{3}\right]_{95}^{103} = 9806.33$

$9806.33 - [99^2] = \frac{16}{3}$

$V(x) = \frac{16}{3}$ $\quad \sqrt{V(x)} = \sigma \quad \sqrt{\frac{16}{3}} = 2.309$

$\boxed{\text{The standard deviation is 2.309}}$

c) $P(x > 100) = 1 - P(x < 100)$

$= 1 - \int_{95}^{100} \frac{1}{8} dx$

$= 1 - 0.625 = 0.375$

$\boxed{\text{The probability that the amount of time is greater than 100 minutes is 0.375}}$

d) Binomial distribution

$b(x; n, p) \rightarrow b(3; 6, 0.375)$

$= \binom{6}{3} 0.375^3 (1-0.375)^{6-3}$

$= 20\,(0.375^3)(0.625^3)$

$= 0.25749 \approx 0.26$

$\boxed{\text{The probability that 3 of 6 classes have duration over 100 min is 0.26}}$

2. 5000 signals sent 0.1% of signals fail

a) Use Poisson distribution (large sample, small probability)

$\lambda = 5000 \times 0.001 = 5$

$P(X>5) = 1 - P(X \le 5)$

$P(X \le 5) = e^{-5}\left(\frac{5^0}{0!} + \frac{5^1}{1!} + \frac{5^2}{2!} + \frac{5^3}{3!} + \frac{5^4}{4!} + \frac{5^5}{5!}\right)$

$= 0.6092$

The probability that fewer than 4,995 messages reach is 0.6092

b) The events are independent - the events from one day should not affect the events of another day

P(fails | one day passed without failure) = P(fails)

$P(X=1) = e^{-5}\frac{5^1}{1!} = 0.03368$

The probability that a signal (one signal) will fail to reach the base the next day is 0.03368

(I'm assuming by "a signa" it means one signal. If the question means the rate of failure, its still 0.1% of signals fail).

c) Median is $F(x) = 0.5$

Use exponential distribution method

$F(x) = 1 - e^{-5x} = 0.5$

$e^{-5x} = 0.5$

$x = 0.13862$

The value of the median time is 0.13862

d) Gamma distribution

$T \sim \Gamma(\alpha, \beta)$ $\alpha = 10$ $\beta = \frac{1}{\lambda} = \frac{1}{5}$

$P(1<T<2) = F(\frac{2}{1/5}, 10) - F(\frac{1}{1/5}, 10)$

$= F(10, 10) - F(5, 10)$

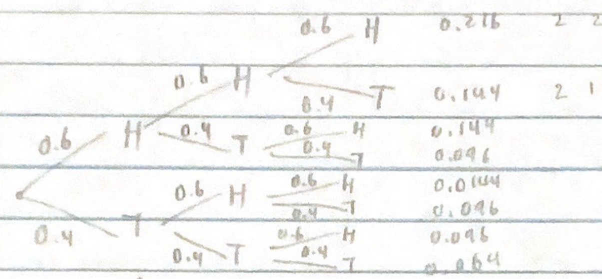
$= 0.542 - 0.032$

$= 0.51$

The probability of this is 0.51

3. heads = 0.6 X : first 2 tosses Y = last two tosses

a)



number of heads

b)

~~X~~	~~0~~	~~1~~	~~2~~	~~$P_X(x)$~~
~~0~~	~~0.064~~	~~0.096~~	~~0~~	
~~1~~	~~0.096~~	~~0.24~~	~~0.216~~	
~~2~~	~~0~~	~~0.144~~	~~0.216~~	
~~$P_Y(y)$~~	~~0.16~~	~~0.384~~	~~0.432~~	

	Y			
X	0	1	2	$P_x(x)$
0	0.064	0.096	0	0.16
1	0.096	0.24	0.144	0.48
2	0	0.144	0.216	0.36
$P_Y(y)$	0.16	0.48	0.36	

c) All possible X given Y = 2

$P(0|2) \rightarrow \frac{P(0 \cap 2)}{P(2)} = \frac{0}{0.36} = 0$

$P(1|2) \rightarrow \frac{P(1 \cap 2)}{P(2)} = \frac{0.144}{0.36} = 0.4$

$P(2|2) \rightarrow \frac{P(2 \cap 2)}{P(2)} = \frac{0.216}{0.36} = 0.6$

$\boxed{P(0|2) = 0 \quad , \quad P(1|2) = 0.4 \quad P(2|2) = 0.6}$

d) Want to show that $p(x,y) = p_X(x) \cdot p_Y(y)$ for all values

$p(0,0) : 0.16 \times 0.16 = 0.064$ X

$p(0,1) : 0.48 \times 0.16 = 0.096$ X

$p(0,2) : 0.16 \times 0.36 = 0$ X

$p(2,0) : 0.16 \times 0.36 = 0$ X

As we can see from these 4 examples, $p(x,y) \neq p_x(x) \cdot p_y(y)$ for all values in the table. Thus, they are not independent of each other.

3e) $\mathrm{corr}(x,y) = \frac{\mathrm{cov}(x,y)}{\sigma_x \sigma_y}$

$E(XY) = \sum_x \sum_y xy\, p(x,y)$

$E(XY) = (1 \times 1 \times 0.24) + (1 \times 2 \times 0.144) + (2 \times 1 \times 0.144) + (2 \times 2 \times 0.216)$

$= 1.68$

$E(X) = \sum_x x \cdot p_x(x) = (0 \times 0.16) + (1 \times 0.48) + (2 \times 0.36) = 1.2$

$E(Y) = \sum_y y \cdot p_y(y) = (0 \times 0.16) + (1 \times 0.48) + (2 \times 0.36) = 1.2$

$\mathrm{cov}(x,y) = E(XY) - E(X)E(Y)$

$\mathrm{cov}(x,y) = 1.68 - 1.2^2 = 0.24$

$V(x) = E(x^2) - [E(x)]^2$

$= (0^2 \times 0.16) + (1^2 \times 0.48) + (2^2 \times 0.36) - 1.44 = 0.48$

$V(y) = E(y^2) - [E(y)]^2$

$= (0^2 \times 0.16) + (1^2 \times 0.48) + (2^2 \times 0.36) - 1.44 = 0.48$

$\sigma_x = \sqrt{V(x)}$ $\quad\quad$ $\sigma_y = \sqrt{V(y)}$

$\sigma_x = 0.693$ $\quad\quad$ $\sigma_y = 0.693$

$\frac{\mathrm{cov}(x,y)}{\sigma_x \sigma_y} = \frac{0.24}{0.693^2} = 0.4997$

$\boxed{\text{correlation coefficient of X and Y is 0.4997} \quad \text{X and Y are correlated}}$

4. mean = 1.86 STD = 0.27 n = 80

a) $P(x < 1.8)$ → I'm assumming normal distribution, CLT

$X \sim N(1.86, 0.27^2)$

$P(x < 1.8)$

$= P\left(Z < \frac{1.8 - 1.86}{0.27}\right)$ -0.22

$= P(Z < -0.22) = 0.4129$

$\boxed{\text{The probababilit that it is less than 1.8 is 0.4129}}$

b) 80^{th} percentile of $X \sim N(1.86, 0.27^2)$

$Z_{80\%} \approx 0.84$

$\frac{x - 1.86}{0.27} = 0.84$

$\boxed{\text{The } 80^{th} \text{ percentile is 2.09}}$

c) $P(x < 1.8) = 0.01$

$X \sim N\left(1.86, \left(\frac{0.27}{\sqrt{n}}\right)^2\right) = 0.01$

$= P\left(Z < \frac{1.8 - 1.86}{\left(\frac{0.27}{\sqrt{n}}\right)}\right) = P(Z < -2.33)$

$n = 109.9$

$n \approx 110$

$\boxed{\text{A sample size of 110 is needed}}$

d) $P(x < 1.8) = 0.01$

$X \sim N\left(1.88, \left(\frac{0.27}{\sqrt{n}}\right)^2\right) = 0.01$

$= P\left(Z < \frac{1.8 - 1.88}{\left(\frac{0.27}{\sqrt{n}}\right)}\right) = P(Z < -2.33)$

$n = 61.83$

$n \approx 62$

$\boxed{\text{A sample size of 62 is needed}}$

5. a) mean: $(6.39 + 6.90 + 7.01 + 6.97 + 6.54 + 6.77 + 6.59 + 6.56 + 6.91 + 6.86) \div 10 = 6.75$

std deviation: $\sqrt{\dfrac{(6.39-6.75)^2 + (6.90-6.75)^2 + \dots (6.86-6.75)^2}{10-1}} = 0.214$

$\boxed{\text{Mean} = 6.75,\ \text{STD} = 0.214}$

b) Null Hypothesis (H_0): mean pH is 7.0 (or more)

Alternative Hypothesis (H_a): mean pH is less than 7.0

c) less than 7.0 → one-tailed confidence interval, t-distribution ($n < 40$)

$H_a: \mu < \mu_0 \rightarrow t \leq -t_{\alpha,n-1}$

$t_{0.05,9} = 1.833 \rightarrow t \leq -1.833$ rejection region

$\boxed{\text{The rejection region is when } t \leq -1.833}$

d) Test Statistic: $t = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \longrightarrow t = \dfrac{6.75 - 7.0}{\frac{0.214}{\sqrt{10}}}$

$t = -3.694$

$-3.694 \leq -1.833$

Reject the null hypothesis because the value of t falls within the rejection region found in part c

e) $\alpha = 0.05$ $t = -3.694$ $v = 9$ [look at t curve tail Areas]

$t \approx -3.7$ $v = 9$ → 0.02

P-value = 0.02, which is less than our $\alpha = 0.05$. Therefore, we reject the null hypothesis, just like in part d

f) $\alpha = 0.05$ $\bar{x} + t_{0.05,9}\left(\frac{0.214}{\sqrt{10}}\right)$

$6.75 + 1.833\left(\frac{0.214}{\sqrt{10}}\right)$

$\mu < 6.874$

This mean pH is less than 7.0, which means we reject the null hypothesis. This is consistent with our previous findings from hypothesis testing and p-value

Final Reference Sheet

Ethan Wong

(PDF) Probability Density Function: $P(a \le X \le b) = \int_a^b f(x)\,dx$

uniformly distributed - $f(x; A, B) = \begin{cases} \frac{1}{B-A} & A \le x \le B \\ 0 & \text{else} \end{cases}$

(CDF) Cumulative Distribution Function: $F(x) = P(X \le x)\ \int_{-\infty}^{x} f(y)\,dy$

$P(X > a) = 1 - F(a)$ $\quad P(X < a) = F(a)$ $\quad P(a \le X \le b) = F(b) - F(a)$ OR $\int_a^b f(x)\,dx$

$cdf = \int_0^x f(x)\,dx$ [f(x) is function from pdf] $\frac{d}{dx} cdf = pdf$

Percentile: $\int_a^x f(x)\,dx = \text{percentile}$ a: lower bound x: solving for x as answer

Expected Value (mean): $E(x) = \mu = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$

Variance: $V(x) = E(x^2) - [E(x)]^2$ OR $V(x) = \int_{-\infty}^{\infty} (x - \underset{\text{EV/mean}}{\mu})^2 \cdot f(x)\,dx$

week 6

Use Z-table

$P(z \le -0.75)$

Go to -0.7 row
Go to 0.05 column

Normal Distribution (Gaussian Distribution) - looks like Bell Curve $X \sim N(\mu, \sigma^2)$

Z: standard normal RV $\quad Z = \frac{X - \mu}{\sigma}$

X: Value you're looking for (i.e $P(x \le \boxed{69})$) $\quad \mu = \text{population mean}$ $\quad \sigma$ = population STD

"30th percentile of standard normal" -> look for value closest to 0.3, use that Z-value

"What's the X^{th} percentile of $X \sim N(a, b^2)$" -> $Z_x \approx ?$ $\quad \frac{x-a}{b} = ?$ solve for x HW5 #1c

Approximating binomial distribution: WK 6 Thursday idk

week 7 Tues

HW5 #3

Lognormal Distribution - if $y = \ln(x)$ has normal dist. with parameters μ and σ (μ: mean of $y = \ln(x)$, σ = STD of $y = \ln(x)$)

- pretty similar to normal distribution but uses $\ln(x)$ instead of x

Expected Value: $E(x) = e^{\mu + \frac{\sigma^2}{2}}$

Variance: $V(x) = e^{2\mu + \sigma^2} \cdot (e^{\sigma^2} - 1)$

week 7 Tues

HW5 #4

Exponential Distribution - probability until next occurrence of an event

$\lambda = \alpha$ (expected number of events to occur in one unit of time)

cdf: $F(t; \lambda) = \begin{cases} 1 - e^{-\lambda t} & t \ge 0 \\ 0 & \text{else} \end{cases}$ (integral of pdf) Mainly use this

pdf: $f(t; \lambda) = \begin{cases} \lambda e^{-\lambda t} & t \ge 0 \\ 0 & \text{else} \end{cases}$ (derivative of cdf)

$P(t \ge 10) = 1 - P(t < 10)$
$= 1 - [1 - e^{-\lambda t}]$

Expected Value / Mean: $E(T) = \frac{1}{\lambda}$ $\quad$ STD $= \frac{1}{\lambda}$ as well

Exponential Distribution is Memoryless - doesn't matter what happened in past

WEEK 7 TUES

Gamma Distribution - $X \sim \Gamma(\alpha, \beta)$ • probability of finding occurrence of event between Time X and Time Y

HW5 #5 $E(x) = \mu = \alpha\beta$ $\overset{\alpha\ \text{integer}}{\beta = \frac{1}{\lambda}}$ $V(x) = \sigma^2 = \alpha\beta^2$

HW5 #6 d   cdf $P(X \le x) = F(x; \alpha, \beta) = F(\frac{x}{\beta}, \alpha)$

week 7 Thurs

(1) Joint probability mass function (pmf) - 2 RVs x, y

$P(x, y) = P(X = x \text{ and } Y = y)$

$p(x,y) \ge 0$, $\sum_x \sum_y p(x,y) = 1$

(2) Marginal pmf of x, y - table of values

$p_X(x) = P(X = x) = \sum_y p(x,y)$

$p_Y(y) = P(Y = y) = \sum_x p(x,y)$

HW6 #5

(3) Joint pdf : $P[X, Y \in A] = \int_c^d \int_a^b f(x,y)\,dx\,dy$

(4) Marginal pdf : $f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$ $f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$

[Replace infinity with actual domain of $f(x,y)$]

Independence of RV: $p(x,y) = p_X(x) \cdot p_Y(y)$ [X, Y are discrete]

$f(x,y) = f_X(x) \cdot f_Y(y)$ [X, Y are continuous]

HW6 #2

Conditional distribution $p_{Y|X}(y|x) = \frac{p(x,y)}{p_X(x)}$ and vice versa [usually given table?]

week 8

Expected Value of X and Y

$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x,y)\,dx\,dy$, $E(XY) = \sum_x \sum_y XY\, p(x,y)$ (usually given table for 2nd version)

Covariance : $Cov(XY) = E(XY) - E(x)E(y)$

Correlation Coefficient : $\rho_{x,y} = Corr(x,y) = \frac{Cov(XY)}{\sigma_x \sigma_y}$

Get σ_x and σ_y by finding $\sqrt{V(x)}$ and $\sqrt{V(y)}$ $V(x) = E(x^2) - [E(x)]^2$

Independence : if $Cov(x,y) = 0$, then x and y are uncorrelated

$\rho_{x,y} = 0$ doesn't necessarily guarantee independence though

Linear Combinations of RVs : idk lol NB