1. With the exception of digital communication channels with additive Gaussian noise, we haven't dealt much with situations involving related continuous and discrete random variables defined on the same probability space. Such situations arise frequently in applications. The key consideration when modeling such situations probabilistically is that your model needs to specify enough information to determine $\mathbb{P}(A)$ for any event involving the random variables. For example, suppose X and Z are discrete random variables and Y is a continuous random variable. Suppose we're given $p_{X,Z}(x,z)$ and $f_Y(y \mid x,z)$ for all x and z for which $p_{X,Z}(x,z) > 0$.

(a) Every event involving these three random variables can be expressed in terms of events A of the form

$$A = \{X = x\} \cap \{Y \in V\} \cap \{Z = z\} \ .$$

Write an expression for $\mathbb{P}(A)$ in terms of the givens above.

(b) Express $p_{X|Z}(x \mid z)$ and $f_{Y|Z}(y \mid z)$ in terms of the givens.

(c) Dealing with conditional probability and conditional expectation in these situations can be tricky, but we can always use a limiting process to work things out. For example, suppose we want $\mathbb{E}(X \mid Y = y, Z = z)$. We can figure this out by first computing $p_{X|Y,Z}(x \mid y,z)$, from which we obtain

$$\mathbb{E}(X \mid Y = y, Z = z) = \sum_x x p_{X|Y,Z}(x \mid y,z) \ .$$

Use a limiting process to derive the (unintuitive to me, at least) formula

$$p_{X|Y,Z}(x \mid y,z) = \frac{f_{Y|X,Z}(y \mid x,z) p_{X|Z}(x \mid z)}{f_{Y|Z}(y \mid z)} \ .$$

2. Discrete and continuous random variables also arise in situations such as the following. Suppose W is a Bernoulli random variable with parameter $q = 2/3$, i.e. $\mathbb{P}(\{W = 1\}) = 2/3$. Let X and Y, independent of each other and of W, be respectively an exponential random variable with rate parameter λ and a geometric random variable with rate parameter p. Finally, define the random variable

$$Z = \begin{cases} X & \text{when } W = 1 \\ Y & \text{when } W = 0 \ . \end{cases}$$

Z is neither a discrete random variable nor a continuous one — we call Z a mixed random variable.

(a) Find $\mathbb{E}(Z \mid W)$. Write it as a one-line formula involving W.

(b) Every random variable, including Z, has a cdf. Find a formula for $F_Z(z)$. You might want to find the formulas for $F_X(x)$ and $F_Y(y)$ first.

(c) Someone on Ho Plaza and presents you with the function

$$F(z) = \begin{cases} 0 & \text{when } z < 0 \\ z/28 & \text{when } 0 \le z < 2 \\ z/28 + 1/4 & \text{when } 2 \le z < 4 \\ z/28 + 1/2 & \text{when } 4 \le z < 6 \\ z/28 + 3/4 & \text{when } 6 \le z < 7 \\ 1 & \text{when } z \ge 7 \ . \end{cases}$$

Find a Bernoulli random variable W', a discrete random variable X', and a continuous random variable Y' so that $F(z)$ is the cdf of the random variable Z'

defined by

$$Z' = \begin{cases} X' & \text{when } W' = 1 \\ Y' & \text{when } W' = 0 \,. \end{cases}$$

3. You can't improve in general on the Markov and Chebyshev inequalities — i.e. for each inequality there exist random variables for which the inequality holds with equality.

(a) Exhibit such a random variable X for the Chebyshev inequality.

(b) Exhibit such a random variable Y for the Markov inequality.

4. From past experience, a professor knows that the test score of a randomly selected student taking her final exam is a continuous random variable X with mean 75.

(a) Find an upper bound on the probability that a student's score will exceed 85. Your upper bound must depend only on the information given so far.

(b) Suppose that the professor knows also that the variance of a randomly selected student's test score is 28. Provide a lower bound on the probability that a student will score between 65 and 85.

(c) How many students would need to take the exam to ensure that, with probability at least 0.9, the class average will be within 7 points of 75? Please don't use the central limit theorem here.

5. Consider a sequence of iid random variables $X_1, \ldots, X_n$ with cdf $F(x)$. You don't know $F(x)$ but would like to estimate it from the sequence $\{X_k\}$. A popular estimator that people use in practice is the *empirical cdf.* For each n and x, define the random variable $G_n(x)$ by

$$G_n(x) = \frac{\text{number of elements in } \{X_k : 1 \le k \le n\} \ \le x}{n} = \frac{1}{n}\sum_{k=1}^{n} \chi_{(-\infty,x]}(X_k) \,,$$

where the so-called indicator function $\chi_{(-\infty,x]}$ defined by the formula

$$\chi_{(-\infty,x]}(X_k) = \begin{cases} 1 & \text{when } X_k \le x \\ 0 & \text{when } X_k > x \,. \end{cases}$$

(a) For fixed values of n and x, what kind of random variable is $nG_n(x)$? What are its mean and variance?

(b) Use Chebyshev's inequality to show that for each value of x, $G_n(x)$ converges in probability to $F(x)$ as $n \to \infty$.

(c) Use the central limit theorem to show that for each value of x the random variable

$$\sqrt{n}\,(G_n(x) - F(x))$$

is asymptotically (as $n \to \infty$) Gaussian with mean 0 and variance $F(x)(1-F(x))$.

Rami Rellumbi

HW12 5/7/19

- rp534 -

① Given $p_{X,Z}(x,z)$ and $f_Y(y|x,z)$ $\forall$ x, z for which $p_{X,Z}(x,z) > 0$

(a) Every event involving X, Y, Z can be expressed in terms of events A of the form

$$A = \{X=x\} \cap \{Y \in V\} \cap \{Z=z\}$$

$\underline{\mathbb{P}(A)\,?}$

$$\mathbb{P}(A) = \mathbb{P}\Big(\{X=x\} \cap \{Y \in V\} \cap \{Z=z\}\Big)$$

$$= \mathbb{P}\Big(\{X=x\} \cap \{Z=z\} \cap \{Y \in V\}\Big)$$

$$= \mathbb{P}\Big(\{Y \in V\} \,\Big|\, \{X=x\} \cap \{Z=z\}\Big)\,\mathbb{P}\Big(\{X=x\} \cap \{Z=z\}\Big)$$

$$= \int_V f_{Y|X,Z}(y|x,z)\, p_{X,Z}(x,z)$$

(b) $p_{X|Z}(x|z)$, $f_{Y|Z}(y|z)$ in term of givens.

$$p_{X|Z}(x|z) = \frac{\mathbb{P}(\{X=x\} \cap \{Z=z\})}{\mathbb{P}(\{Z=z\})} = \frac{p_{X,Z}(x,z)}{\sum_t p_{X,Z}(t,z)}$$

$$f_{Y|Z}(y|z) = \frac{\mathbb{P}(\{Y \in V\} \cap \{Z=z\})}{\mathbb{P}(\{Z=z\})} = \frac{\sum_x f_{Y|X,Z}}{\sum_x p_{X,Z}(x,z)}$$

(c) Use limiting process to derive

$$P_{X|Y,Z}(x|y,z) = \frac{f_{Y|X,Z}(y|x,z)\, p_{X|Z}(x|z)}{f_{Y|Z}(y|z)}$$

$$P_{X|Y,Z}(x|y,z) = \lim_{\delta \to 0} \mathbb{P}\left(\{X=x\} \mid \{Y \in [y-\delta, y+\delta]\} \cap \{Z=z\}\right)$$

$$= \lim_{\delta \to 0} \frac{\mathbb{P}\left(\{X=x\} \cap \{Y \in [y-\delta, y+\delta]\} \cap \{Z=z\}\right)}{\mathbb{P}\left(\{Y \in [y-\delta, y+\delta]\} \cap \{Z=z\}\right)}$$

$$= \lim_{\delta \to 0} \frac{\mathbb{P}\left(\{Y \in [y-\delta, y+\delta]\} \mid \{X=x\} \cap \{Z=z\}\right) \mathbb{P}\left(\{X=x\} \cap \{Z=z\}\right)}{\mathbb{P}\left(\{Y \in [y-\delta, y+\delta]\} \mid \{Z=z\}\right) \mathbb{P}\left(\{Z=z\}\right)}$$

$$= \lim_{\delta \to 0} \frac{\mathbb{P}\left(\{Y \in [y-\delta, y+\delta]\} \mid \{X=x\} \cap \{Z=z\}\right) p_{X,Z}(x,z)}{\mathbb{P}\left(\{Y \in [y-\delta, y+\delta]\} \mid \{Z=z\}\right) p_Z(z)}$$

$$= \lim_{\delta \to 0} \frac{\left(\int_{y-\delta}^{y+\delta} f_{Y|X,Z}(y|x,z)\, dy\right) p_{X,Z}(x,z)}{\left(\int_{y-\delta}^{y+\delta} f_{Y|Z}(y|z)\, dy\right) p_Z(z)}$$

$$= \lim_{\delta \to 0} \frac{\frac{1}{2\delta}\left(\int_{y-\delta}^{y+\delta} f_{Y|X,Z}(y|x,z)\,dy\right) p_{X,Z}(x,z)}{\frac{1}{2\delta}\left(\int_{y-\delta}^{y+\delta} f_{Y|Z}(y|z)\,dy\right) p_Z(z)}$$

$$= \frac{f_{Y|X,Z}(y|x,z)\; p_{X|Z}(x|z)}{f_{Y|Z}(y|z)}$$

② $W \sim \text{Bernoulli}(q)$; $q = 2/3$ ie $\mathbb{P}(\{W=1\}) = 2/3$

X independent of Y; both independent of W

$X \sim \text{Exponential}(\lambda)$

$Y \sim \text{Geometric}(p)$

Define

Z is neither continuous or discrete; say it's mixed

$$Z = \begin{cases} Y, & \text{when } W=0 \\ X, & \text{when } W=1 \end{cases}$$

(a) $\mathbb{E}(Z \mid W)$?

$\mathbb{E}(Z \mid W=0) = \mathbb{E}(Y) = 1/p$

$\mathbb{E}(Z \mid W=1) = \mathbb{E}(X) = 1/\lambda$

So

$$\mathbb{E}(Z \mid W=w) = \begin{cases} \mathbb{E}(Y), & w=0 \\ \mathbb{E}(X), & w=1 \end{cases} = w\left(\frac{1}{\lambda}\right) + (1-w)\frac{1}{p}$$

Thus

$$\mathbb{E}(Z \mid W) = \frac{W}{\lambda} + \frac{(1-W)}{p}$$

EXTRA I DIDNT WANT TO ERASE

$$\mathbb{E}(Z) = \sum_{w \in W} \mathbb{E}(Z \mid W=w)\, \mathbb{P}(\{W=w\})$$

$$= \mathbb{E}(Z \mid W=0)\, \mathbb{P}(\{W=0\}) + \mathbb{E}(Z \mid W=1)\, \mathbb{P}(\{W=1\})$$

$$= \frac{1}{p}\cdot\frac{1}{3} + \frac{1}{\lambda}\cdot\frac{2}{3}$$

(b) $F_Z(z)$? (All random variables have a cdf!)

$$F_X(x) = \mathbb{P}(\{X \le x\}) = \begin{cases} \int_0^x \lambda e^{-\lambda x} dx & , x \geq 0 \\ 0 & \text{else} \end{cases} = \begin{cases} 1 - e^{-\lambda x} & , x \geq 0 \\ 0 & , \text{else} \end{cases}$$

$$F_Y(y) = \mathbb{P}(Y \le y) = 1 - (1-p)^y$$

$$\begin{aligned} F_Z(z) &= \mathbb{P}(Z \le z) \\ &= \mathbb{P}(\{Z \le z\} \mid \{W = 1\}) \mathbb{P}(\{W=1\}) + \mathbb{P}(\{Z \le z\} \mid \{W=0\}) \mathbb{P}(\{W=0\}) \\ &= 2/3 \, \mathbb{P}(\{X \le z\}) + \tfrac{1}{3} \mathbb{P}(\{Y \le z\}) \\ &= \tfrac{2}{3} F_X(z) + \tfrac{1}{3} F_Y(z) \end{aligned}$$

$$F_Z(z) = \begin{cases} 0 & , z < 0 \\ \frac{2}{3}(1 - e^{-\lambda z}) & , 0 \le z < 1 \\ \frac{2}{3}(1 - e^{-\lambda z}) + \frac{1}{3}(1 - (1-p)^k) & , k \le z < k+1 \, , k \geq 1 \end{cases}$$

(c)

$$F_Z(z) = \begin{cases} 0 & ; \ z < 0 \\ z/28 & ; \ 0 \le z < 2 \\ z/28 + 1/4 & ; \ 2 \le z < 4 \\ z/28 + 1/2 & ; \ 4 \le z < 6 \\ z/28 + 3/4 & ; \ 6 \le z < 7 \\ 1 & ; \ z \geq 7 \end{cases}$$

Find $W' \sim \text{Bernoulli}(q')$; X' discrete, Y' continuous so $F(z)$ is the cdf of Z'.

$F_Z(z)$ has jumps at $z = 2, 4, 6$.

- Expect discrete X' to take on positive probability values here.
- X' will have staircase cdf.
  - jumps of same size $\Rightarrow$ X' uniform pmf over those three values $\rightarrow$ probability $1/3$ for each.
- Need $p_{W'}(1) = 3/4$ to make jumps in $F(z)$ right size

Thus need

$$\frac{1}{4} F_{Y'}(z) = \begin{cases} 0 & , \ z < 0 \\ z/28 & , \ 0 \le z < 7 \\ 1 & , \ z \geq 7 \end{cases} \quad \leftarrow Y' \sim \text{Uniform}[0,7]$$

To recap

$W' \sim \text{Bernoulli}(3/4)$ i.e $\mathbb{P}(W=1) = 3/4$; $\mathbb{P}(W=0) = 1/4$

X' discrete with $p_{X'}(2) = p_{X'}(4) = p_{X'}(6) = 1/3$

$Y' \sim \text{Uniform}[0,7]$

(3)

(a) A rv X for which the Chebyshev inequality holds w/ equality.

Chebyshev Inequality: $\mathbb{P}(|X-\mu| \geq c) \leq \frac{Var(X)}{c^2}$

Want X such that

$$\mathbb{P}(|X-\mu| \geq c) = \frac{Var(X)}{c^2}$$

Define

$$X = \begin{cases} -1, & \text{w/ probability } 1/2k^2 \\ 0, & \text{w/ probability } 1-1/k^2 \\ 1, & \text{w/ probability } 1/2k^2 \end{cases} \qquad \forall\, k \geq 1$$

Then

$$\mathbb{E}(X) = \frac{-1}{2k^2} + \frac{1}{2k^2} + 0 = 0$$

$$Var(X) = \mathbb{E}(X^2) = \frac{1}{2k^2} + \frac{1}{2k^2} = \frac{1}{k^2} \;;\; \sigma^2 = \frac{1}{k^2},\ \sigma = \frac{1}{k}$$

Thus

$$\mathbb{P}(|X-\mu| \geq k\sigma) = \mathbb{P}(|X| \geq 1) = \frac{1}{k^2}$$

This works for any linear transformation as well.

(b) Markov Inequality: $\mathbb{P}(\{X \geq c\}) \leq \frac{E(X)}{c}$

Want X such that

$$\mathbb{P}(X \geq c) = \frac{\mathbb{E}(X)}{c}$$

If you think about it, the Chebyshev Inequality follows from the Markov Inequality (for rv $[X - \mathbb{E}(X)]^2$)!
So the condition for equality is the same.

The above were realized by following the proof of Markov/Chebyshev inequality on Wikipedia.

④ $X \sim$ Continuous rv indicating test score

$\mathbb{E}(X) = 75$

(a) Upper bound a student's score will exceed 85?

$$\mathbb{P}(X>85) < \frac{\mathbb{E}(X)}{85} = \frac{75}{85} = \frac{15}{17} \cong 0.88$$

(b) $Var(X) = 28$.

Lower bound on probability a student scores between 65 and 85.

$$\mathbb{P}(65 < X < 85) = \mathbb{P}(|X-75| < 10)$$
$$= 1 - \mathbb{P}(|X-75| \geq 10)$$
$$> 1 - \frac{Var(X)}{100}$$
$$> 1 - \frac{28}{100}$$
$$\boxed{> 0.72}$$

(c) How many students would have to take exam to ensure that, w/ probability at least 0.9, the class average will be within 7 points of 75?

Each student has a test score X. Each X is iid.

Want amount of students $-n-$ such that class average is within 7 points of 75.

Let $\{X_n\}$ be a sequence of student test scores.

Let $M_n = \frac{X_1 + \cdots + X_n}{n}$

i.e Want

$$\mathbb{P}(|M_n - 75| < 7) > 0.9 \qquad (\mathbb{E}(X) = 75)$$

$$[1 - \mathbb{P}(|M_n - 75| \geq 7)] > 0.9$$

$$\mathbb{P}(|M_n - 75| \geq 7) < 0.1$$

So

$$\mathbb{P}(|M_n - 75 \geq 7) \leq \frac{Var(X)}{7^2 n} < 0.1$$

$$\Rightarrow \frac{28}{7^2 n} < 0.1$$

Thus

$$n > \frac{28}{7^2(0.1)} = 5.7 \rightarrow$$ At least 6 students must take exam

(5) • Sequence of iid random variables $X_1, \dots, X_n$ with cdf $F(x)$.

• Know $F(x)$; want to estimate it from the sequence $\{X_k\}$

$\forall n, x$ define the random variable

$$G_n(x) = \frac{(\text{number of elements in } \{X_k : 1 \le k \le n\}) \le x}{n} = \frac{1}{n}\sum_{k=1}^{n} \chi_{(-\infty, x]}(X_k)$$

where

$$\chi_{(-\infty,x]}(X_k) = \begin{cases} 1 ; & X_k \le x \\ 0 ; & X_k > x \end{cases}$$

(a) For fixed n, x; $G_n(x)$ is a Binomial random variable.

To see this, observe

$$G_n(x) = \frac{\text{Elements in set} \le \text{Constant}_1}{\text{Constant}_2} \leftarrow \text{normalizes it}$$

i.e

$$0 \le \frac{1}{n}\sum_{k=1}^{n} \chi_{(-\infty,x]}(X_k) \le 1$$

Each X_k is either $\le x$ or it isn't. So G_n is the sum of n independent Bernoulli random variables. This random variable has parameter $p = \mathbb{P}(\{X \le x\}) = F(x)$. The mean is thus $F(x)$ and variance is $F(x)(1-F(x))$

(b) Using Chebyshev,

$$\mathbb{P}(|X-\mu| \geq c) \leq \frac{Var(x)}{c^2}$$

Fix $\varepsilon > 0$, want

$$\lim_{n \to \infty} \mathbb{P}(| G_n(x) - F(x) | \geq \varepsilon) = 0$$

$$\mathbb{P}(| G_n(x) - F(x) | \geq \varepsilon) \leq \frac{F(x)(1-F(x))}{\varepsilon^2 n}$$

Obvious to see this $\to 0$ as $n \to \infty$

i.e $G_n(x)$ converges in probability to $F(x)$ as $n \to \infty$

(c) $\sqrt{n}\left(G_n(x) - F(x)\right) \xrightarrow{n} \text{Gaussian}\left(0, F(x)(1-F(x))\right)$

Since $\forall x$ $G_n(x)$ is $\frac{1}{n}$ times the sum of n iid random variables with common mean $F(x)$ and common variance $F(x)(1-F(x))$, Central Limit Theorem says that

μ

$$W_n = \frac{\sqrt{n}}{\sqrt{F(x)(1-F(x))}}\left(G_n(x) - F(x)\right)$$

σ

is asymptotically Gaussian with mean zero, variance 1.

Since

$$\sqrt{n}\left(G_n(x - F(x))\right) = \sqrt{F(x)(1-F(x))}\, W_n$$

This is a Gaussian with mean 0, Variance $F(x)(1-F(x))$.

To see this, observe

$$\mathbb{E}\left(\sqrt{F(x)(1-F(x))}\, W_n\right) = \sqrt{F(x)(1-F(x))}\, \mathbb{E}(W_n) = 0$$

$$\mathrm{Var}\left(\sqrt{F(x)(1-F(x))}\, W_n\right) = \left(\sqrt{F(x)(1-F(x))}\right)^2 \mathrm{Var}(W_n) = F(x)(1-F(x))$$