Probability Review

Klein
carlj.klein@gmail.com

1 Cards

Question1:
Inclusion-Exclusion ID
Hint: P(A $\bigcup$ B) = ?
Answer1:
P(A $\bigcup$ B) = P(A) + P(B) - P(AB)

Question2:
Define Mutually Exclusive
Answer2:
If AB = $\emptyset \rightarrowtail$ A and B are mutually exclusive

Question3:
Conditional Probability and Corollary
Answer3:
Definition: P(E|F) = $\frac{P(EF)}{P(F)}$
Corollary: P(EF) = P(E) * P(F|E)

Question4:
Multiplication Rule
Hint: extension of conditional probability: $\rightarrow P(E_1 *** E_n) = ?$
Answer4:
$P(E_1 *** E_n) = P(E_1) * P(E_2 \mid E_1) * P(E_3 \mid E_2 * E_1) *** P(E_n \mid E_{n-1} *** E_1)$

Question5:
Law of Total Probability
Hint: Given a mutually exclusive and exhaustive set A, P(A_1) + ... + P(A_k) = 1, what can be deduced about the probability of an event B occurring?
Answer5:
P(B)

= P(BA_1) + ... + P(BA_k)
= P(A_1) * P(B | A_1) + ... + P(A_k) * P(B | A_k)
= $\sum_{i=1}^{k} P(A_i) * P(B \mid A_i)$

Question6:
Bayes' Theorem
Answer6:
Given a mutually exclusive and exhaustive set A, P(A_1) + ... + P(A_k) = 1, then
$P(A_j \mid B) =$
$\frac{P(A_j B)}{P(B)} =$
$\frac{P(A_j) * P(B|A_j)}{\sum_{i=1}^{k} P(A_i) * P(B|A_i)}$

Question7:
How does independence extend to intersections and conditional probability?
Answer7:
Given independent events, A and B:
P(AB) = P(A) * P(B)
P(A|B) = P(A) and P(B|A) = P(B)

Question8:
Cumulative Distribution Function (CDF)
Answer8:
F(x) = P(X $\leq$ x)

Question9:
Properties of the CDF
Answer9:

- 0 $\leq$ F(x) $\leq$ 1
- If x $\leq$ y $\rightarrow$ F(x) $\leq$ F(y)
- $lim_{x \rightarrow \infty} F(x) = 1$
- $lim_{xto - \infty} F(x) = 0$

Question10:
Probability Mass Function (pmf)
Answer10:
Note: for finite or countably infinite set (i.e. discrete)
p(x) = P(X=x)

Question11:
For discrete values, describe the following:

- E[X]
- E[X^2]
- E[g(X)]
- E[aX + b]
- E[aX + bY]

Answer11:

- E[X] = $\sum_x xP(X = x)$
- E[X^2] = $\sum_x x^2 P(X = x)$
- E[g(X)] = $\sum_x g(x)P(X = x)$
- E[aX + b] = aE[X] + b
- E[aX + bY] = aE[X] + bE[Y]

Question12:
For discrete values, describe the following:

- Var(X)
- Var(cX)
- Var(aX + b)
- sd(X)

Answer12:

- Var(X) = E[(X - μ)^2] = E[X^2] - (E[X])^2 = σ^2
- Var(cX) = c^{2}Var(X)
- Var(aX + b) = a^{2}Var(X)
- sd(X) = $\sqrt{Var(X)} = \sigma$

Question13:
Define moment and compare it to σ

Answer13:
E[X^k] = k^{th} raw moment
σ = mean = 1st raw moment

Question14:
Define covariance and provide the formula.
Answer14:
Cov(X, Y) =
E[(X - E[X])(Y - E[Y])] =
E[XY] - E[X]E[Y]

Question15:
Describe the following properties of covariance:

- Cov(X, X)
- Cov(aX, Y)
- Cov($\sum_{i=1}^{n} X_i$, $\sum_{j=1}^{m} Y_j$)

Answer15:

- Cov(X, X) = E[X^2] - $(E[X])^2$ = σ^2 = Var(X)
- Cov(aX, Y) = aCov(X, Y)
- Cov($\sum_{i=1}^{n} X_i$, $\sum_{j=1}^{m} Y_j$) = $\sum_{i=1}^{n} \sum_{j=1}^{m}$ Cov(X_i, Y_j)

Question16:
Uniform Distribution (discrete)
Provide the following:

- pmf
- E[X]
- Var(X)

Answer16:

- P(X=x) = $\frac{1}{n}$
- E[X] = $\frac{n+1}{2}$ = $\frac{b+a}{2}$

- Var(X) = $\frac{n^2+1}{12}$ = $\frac{(b-a+1)^2+1}{12}$

Question17:
Bernoulli Distribution (discrete)
Provide the following:

- pmf
- E[X]
- Var(X)

Answer17:
Example: flipping a coin for a single experiment, 0 represents heads (probability p) and 1 represents tails ()probability 1 - p).

- $$P(X = x) = \begin{cases} p; x = 0 \\ 1 - p; x = 1 \end{cases}$$
- E[X] = p
- Var(X) = p(1 - p)

Question18:
Binomial Distribution (discrete)
Provide the following:

- pmf
- E[X]
- Var(X)

Answer18:
Example: Multiple experiments where each experiment only has 2 outcomes (flipping a coin multiple times).
The distribution is the sum of Bernoulli's, number of successes with probability p in n independent trials.

- P(X=x) = $\binom{n}{x}p^x p(1-p)^{n-x}$
- E[X] = np
- Var(X) = np(1 - p)

Question19:
Geometric Distribution - "Starting at 0" (discrete)
Provide the following:

- pmf

- E[X]
- Var(X)

Answer19:
Example: x is the number of failures before a success (i.e. can be 0).

- P(X=x) = p(1-p)^x
- E[X] = $\frac{1-p}{p}$
- Var(X) = $\frac{1-p}{p^2}$

Question20:
Geometric Distribution - "Starting at 1" (discrete)
Provide the following:

- pmf
- E[X]
- Var(X)

Answer20:
Example: x is the number of trials necessary for a success, including the success (i.e. cannot be 0).

- P(X=x) = p(1-p)$^{x-1}$
- E[X] = $\frac{1}{p}$
- Var(X) = $\frac{1-p}{p^2}$

Question21:
Negative Binomial (discrete)
Provide the following:

- pmf
- E[X]
- Var(X)

Answer21:
Number of failures before the rth success.
Sum of Geometric Distributions "starting at 0".

- P(X=x) = $\binom{n+(r-1)}{r-1}p^r(1-p)^n$
- E[X] = $r\frac{1-p}{p}$
- Var(X) = $r\frac{1-p}{p^2}$

Question22:
Poisson Distribution (discrete)
Provide the following:

- pmf
- E[X]
- Var(X)

Answer22:

- P(X=x) = $e^{-\lambda}\frac{\lambda^x}{x!}$
- E[X] = λ
- Var(X) = λ

Note that the sum of independent Poisson distributions is Poisson:
Given N = Poiss(λ_1) and M = Poiss(λ_2)
$\rightarrow$ P(N + M = x) = P(Poiss($\lambda_1 + \lambda_2$) = x)

Question23:
Hypergeometric Distribution (discrete)
Provide the following:

- pmf
- E[X]
- Var(X)

Answer23:
Choosing items without replacement:

- G: total "good"
- g: number of "good" wanted
- N: total
- n: number inspected
- P(X=g) = $\frac{\binom{G}{g}\binom{N-G}{n-g}}{\binom{N}{n}}$
- E[X] = $\frac{G}{N}$
- Var(X) = $\frac{N-n}{N-1}np(1-p)$

Question24:
Multinomial Distribution (discrete)
Provide the following:

- pmf
- E[X]
- Var(X)

Answer24:
Example: k-sided die rolled n times (i.e. n trials with k outcomes). Derivation of the binomial distribution (i.e. has more than 2 outcomes).
Given:
X_i: number of trials resulting in the i^{th} outcome.
k: number of categories
then

- P($X_1 = x_1$, ..., $X_k = x_k$) = $\frac{n!}{x_1!***x_k!} p_1^{x_1} *** p_k^{x_k}$
- E[X_i] $= np_i$
- Var(X_i) $= np_i(1 - p_i)$

Question25:
Define the probability density function (PDF).
Answer25:
Function which describes the probability for an event to occur for continuous functions.

f(x) = $\frac{d}{dx}$F(x) $\rightarrow$ P(X $\in$ B) = $\int_B$f(x)dx

Question26:
Properties of the PDF.
Answer26:
f(x) $\geq$ 0
$\int_{-\infty}^{\infty}$f(x)dx = 1
$\int_a^b$f(x)dx = P(a $\leq$ x $\leq$ b)

Question27:
For continuous values, describe the following:

- E[X]
- E[g(X)]

- Var(X)

Answer27:

- E[X] = $\int_{-\infty}^{\infty}$xf(x)dx
- E[g(X)] = $\int_{-\infty}^{\infty}$g(x)f(x)dx
- Var(X) = E[X^2] - $(E[X])^2$ = σ^2

Question28:
Uniform Distribution (continuous)
Provide the following:

- PDF
- E[X]
- Var(X)

Answer28:

- f(x) = $\frac{1}{b-a}$, $a < x < b$
- E[X] = $\frac{a+b}{2}$
- Var(X) = $\frac{(b-a)^2}{12}$

Question29:
Normal Distribution (continuous)
Provide the following:

- PDF
- E[X]
- Var(X)

Answer29:

- f(x) = $\frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-(x-\mu)^2}{2\sigma^2}}$
- E[X] = μ
- Var(X) = σ^2

Standard Normal Conversion:
Want $\mu = 0$ and $\sigma = 1$
If X = Norm(μ, σ^2) $\rightarrow$ Z = $\frac{x-\mu}{\sigma}$ is the standard normal.

Question30:
Exponential Distribution (continuous)
Provide the following:

- PDF
- E[X]
- Var(X)

Answer30:
Example: X is the waiting time until an event occurs, when events are always occurring at a random rate of $\mu > 0$.

- f(x) = $\mu e^{-\lambda x}$, $x > 0$
- E[X] = $\frac{1}{\lambda}$
- Var(X) =$\frac{1}{\lambda^2}$

$P(X > x) = e^{-\lambda x}$
X has the memoryless property: implies the probability of an event occurring in the future is independent of how long it has not occurred in the past.
Question31:
Central Limit Theorem (CLT).
Answer31:
The sum of a large number of independent random variables has a distribution that is approximately normal.
That is, given a sequence of individually and identically distributed random variables (iid's), each with mean μ and variance σ^2:
$\frac{x_1+...+x_n-n\mu}{\sigma\sqrt{n}}$ is approximately N(0,1) as n $\rightarrow \infty$

$S_n = x_1 + ... + x_n$, E[S_n] = nμ, SD(S_n) = $\sigma\sqrt{n}$, VarSD(S_n) = σ^2n

$\rightarrow$ $\frac{S_n-n\mu}{\sigma\sqrt{n}}$ is approximately N(0,1).

Question32:
Central Limit Theorem (CLT) Extensions, describe the following:

- Given $S_n = x_1 + ... + x_n$, provide the CDF P($S_n \leq$ x)
- Averages

Answer32:

- P($S_n \leq$ x) = P($\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq \frac{x - n\mu}{\sigma\sqrt{n}}$) = $\phi(\frac{x-n\mu}{\sigma\sqrt{n}})$
- $S_n = \sum x_i \rightarrow \bar{x} = \frac{S}{n} \rightarrow E[\bar{X}] = E[X]$
  $Var(\bar{X}) = \frac{Var(X)}{n}$
  $\rightarrow$ X is approximately $N(E[X], \frac{Var(X)}{n})$

Question33:
Describe the strong law of large numbers.
Answer33:
The average of a sequence of iid's random variables will converge to the mean of that distribution.
Given E[X_i] = μ, $S_n = x_1 + ... + x_n$,
then $\frac{S_n}{n} \rightarrow \mu$ as $n \rightarrow \infty$

Question34:
Moment Generating Function (MGF)
Answer34:
M(t) = E[e^{tx}], which can be generated for both discrete and continuous variables.
M'(0) = E[X] = 1^{st} moment
.
.
.
M^k = E[X^k] = k^{th} moment

Question35:
Define joint CDFs for both discrete (pmf) and continuous.
Answer35:
Discrete: joint pmf: p(i, j) = P(X=i, Y=j)
Individual pmfs from joint: P(X=i) = $\sum_j$ P(i, j), P(X=j) = $\sum_i$ P(i, j)

Continuous: F(x, y) = P(X $\leq$ x, Y $\leq$ y); $-\infty < x, y < \infty$
Individual CDFs (marginal distributions):
$F_x(x) = \lim_{y\rightarrow\infty} F(x, y)$, $F_y(x) = \lim_{x\rightarrow\infty} F(x, y)$

Question36:
Define joint PDFs.
Answer36:
Suppose we have joint PDF f(x, y) such that
P((x, y) $\in$ C) = $\int\int_C$ f(x, y)dxdy

$\rightarrow f_x(x) = \int_{\infty}^{-\infty} f(x,y)dy,\ f_y(y) = \int_{\infty}^{-\infty} f(x,y)dx$

$\rightarrow$ f(x, y) = $\frac{\partial^2}{\partial x \partial y}$ F(x, y)

Question37:

Joint PDF Independence.

Answer37:

Random variables X and Y are independent if for all sets A and B:

P(x ∈ A, y ∈ B) = P(X ∈ A)P(y ∈ B)

Independence creates the following:

- F(a, b) = $F_X(a)F_Y(b)$
- p(x, y) = $p_x(x)p_y(y)$
- f(x, y) = $f_X(x)f_Y(y)$

Question38:

Define joint conditional distributions for both discrete and continuous random variables.

Answer38:

Discrete:

P(Y = y | X = x) = $\frac{P(X=x,Y=y)}{P(X=x)} = \frac{P(Y=x,Y=y)}{\sum_y P(X=x,Y=y)}$

Continuous:

$f_{x|y}(y \mid X = x) = \frac{f(x,y)}{f_x(x)} = \frac{f(x,y)}{\int f(x,y)dy}$

Question39:

Define joint moments:

- E[g(x, y)]
- E[y | X=x]
- E[g(y) | X=x]
- E[Y^k]

Answer39:

- E[g(x, y)] = $\int \int g(x,y)f(x,y)dxdy$
- E[y | X=x] = E[Y | X] = $\int y f_{y|x}(y \mid X = x)$
- E[g(y) | X=x] = $\int g(y) f_{y|x}(y \mid X = x)$
- E[Y^k] = E[E[$Y^k \mid X$]]

Question40:

Define the following in terms of X and Y:

- Var(Y)
- Var($Y \mid X = x$)

Answer40:

- Var(Y) = E[Var(Y | X)] + Var(E[Y | X])
- Var($Y \mid X = x$) = E[$Y^2 \mid X = x$] - $(E[Y \mid X = x])^2$

  = $\int y^2 f_{y|x}(y \mid X = x)dy$ - $(\int y f_{y|x}(y \mid X = x)dy)^2$

Question41:
Define the following covariance:

- Cov(X, Y)
- Var(X + Y)
- Var(aX + bY)
- Cov(X, b)
- Cov(X, X)
- Cov(aX + bY, cZ)

Answer41:

- Cov(X, Y) = E[XY] - E[X]E[Y]
- Var(X + Y) = Var(x) + 2Cov(X, Y) + Var(Y)
- Var(aX + bY) = a^2(x) + 2abCov(X, Y) + b^2Var(Y)
- Cov(X, b) = 0
- Cov(X, X) = Var(X)
- Cov(aX + bY, cZ) = Cov(aX, cZ) + Cov(bY, cZ)
  = acCov(X, Z) + bcCov(Y, Z)

Question42:
Define the multi-variate moment generating functions (MGFs).
Answer42:
$M_{X,Y}(s,t)$ = E[e^{sx+ty}]

$\frac{\partial}{\partial s} M_{x,y} = E[Xe^{sx+ty}]$

$\frac{\partial}{\partial t} M_{x,y} = E[Ye^{sx+ty}]$

$\frac{\partial}{\partial s} M_{x,y} \mid_{0,0} = E[X]$

$\frac{\partial}{\partial t} M_{x,y} \mid_{0,0} = E[Y]$

$\frac{\partial^2}{\partial s \partial t} M_{x,y} \mid_{0,0} = E[XY]$