STAT 150 HOMEWORK #1

FALL 2023

Due Friday, Sep 1st, at 11:59 PM on Gradescope.

1. Pinsky and Karlin, Problem 1.3.11

X and Y are geometric with p(k) = $\pi^k(1-\pi)$ Where $U = min(X,Y)$ and $V = max(X,Y)$ and $W = V - U$ Determine the joint pmf of W and U and show they are independent.

$P(X = i) = \pi^i * (1-\pi)$
$P(X > i) = \sum_{k=i+1}^{\infty} \pi^k(1-\pi) = (1-\pi)\frac{\pi^{i+1}}{1-\pi} = \pi^{i+1}$
$P(U = i) = P(X > i, Y = i) + P(X = i, Y > i) + P(X = i, Y = i)$
$P(U = i) = \pi^i(1-\pi)\pi^{i+1} + \pi^i(1-\pi)\pi^{i+1} + \pi^{2i}(1-\pi)^2$
$P(U = i) = 2((1-\pi) * \pi^{2i+1}) + \pi^{2i}(1-\pi)^2$
$P(U = i) = \pi^{2i}[2((1-\pi) * \pi) + (1-\pi)^2]$
$P(U = i) = \pi^{2i}[2(\pi - \pi^2) + (1-\pi)^2]$
$P(U = i) = \pi^{2i}[2\pi - 2\pi^2 + 1 - 2\pi + \pi^2]$
$P(U = i) = \pi^{2i}[1 - \pi^2]$

$W = \mid X - Y \mid$ so there are three situations. $X > Y, X = Y, X < Y$.
This gives us: $P(W = j) = \sum_{i=0}^{\infty} P(X = i, Y = i + j) + P(X = i + j, Y = i)$.
$P(W = j) = \sum_{i=0}^{\infty} 2(\pi)^{2i+j}(1-\pi)^2$
$P(W = j) = 2(1-\pi)^2(\pi)^j \frac{1}{1-\pi^2}$
$P(W = j) = \frac{2(1-\pi)^2(\pi)^j}{1-\pi^2}$

For independence, we check:
$P(U = i, W = j) = P(U = i, V = j + i) = P(X = i + j, Y = i) + P(X = i, Y = i + j) =$
$= \boxed{2\pi^{2i+j}(1-\pi)^2}$ due to the independence of
$P(U = i) * P(W = j) = \pi^{2i}[1-\pi^2] * \frac{2(1-\pi)^2(\pi)^j}{1-\pi^2} =$
$= \boxed{2\pi^{2i+j}(1-\pi)^2}$

This time around, we have fixed i, and we can see that $P(U = i, W = j) = P(U = i)P(W = j)$ by comparing the equations from above. This makes sense, because by fixing i, the minimum, we have not changed the distribution of the maximum. So the distribution of W, the distance, is independent of the distribution of U, the minimum, when no information about the maximum is given.

2. Pinsky and Karlin, Problem 1.4.3

Let X and Y be independent random variables uniformly distributed over the interval $[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$ for some fixed θ.

Show that $W = X - Y$ has a distribution that is independent of θ with density function:

$$f_W(w) = \begin{cases} 1 + w & \text{for } -1 \leq w \leq 0 \\ 1 - w & \text{for } 0 \leq w \leq 1 \\ 0 & \text{for} |w| > 1 \end{cases}$$

$\boxed{\text{The last case is clear: if } |w| > 1\text{, there is no chance of } X - Y = 0\text{, so the density is zero.}}$

$$f_X(x) = \begin{cases} 1, & \theta - \frac{1}{2} \leq x \leq \theta + \frac{1}{2} \\ 0, & \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} 1, & \theta - \frac{1}{2} \leq y \leq \theta + \frac{1}{2} \\ 0, & \text{otherwise} \end{cases}$$

The convolution of f_X and $-f_Y$ (negative because we're considering subtraction) is:

$$f_W(w) = \int_{-\infty}^{\infty} f_X(x) f_Y(x - w)\, dx$$

Given the support of the distributions, the integral can be restricted to the domain $[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$. However, since the convolution depends on the overlapping regions of f_X and $-f_Y$, the integral can break down into two cases based on the values of w:

It is natural to consider these two cases, as each involves one RV being greater than the other. Here I use 'logic' to show that $(-1 \leq w \leq 1)$ I will not prove this, as I am tired.

The intuition in this problem is what makes it more simple than how we covered in OH although this decreases the level of rigor of my solution, I think it is far more suitable to understand what the convolution is doing.

Sliding the $-f_Y$ from the right to the left of f_X gives us an increasing intersection, giving us (a), up to the point where $w = 0$. Then, the size of the intersection begins to fall, as w goes to 1, giving us (b)

When $-1 \leq w \leq 0$:

$$f_W(w) = \int_{\theta + \frac{1}{2} + w}^{\theta + \frac{1}{2}} dx = w + 1$$

When $0 < w \leq 1$:

$$f_W(w) = \int_{\theta - \frac{1}{2}}^{\theta + \frac{1}{2} - w} dx = 1 - w$$

.

Elsewhere, $f_W(w) = 0$. This is because outside these intervals, the densities do not overlap. Combining the cases, the pdf of W is given as the above, with no θ in the final result, implying that the distribution of W is independent of θ.

3. Pinsky and Karlin, Problem 1.5.7

Let $X_1, X_2, ..., X_n$ be independent random variables that are exponentially distributed with respective parameters $\lambda_1 ... \lambda_n$. Identify the distribution of the minimum $V = min\{X_1, X_2, ..., X_n\}$. Hint: For any real number v, the event $V > v$ is equivalent to $X_1 > v, X_2 > v, ..., X_n > v$.

This is a common question, from memory I know that either the min or max of exponential RV's is also exponential, this has to do with the memoryless property.

$S_v(v) = P(V > v) = P(X_1 > v, X_2 > v, ..., X_n > v) = \Pi_{i=1}^{n}\{X_i > v\} = \Pi_{i=1}^{n}(\lambda_i e^{-\lambda_i v})$

Here we will define λ^* to denote: $\Pi_{i=1}^{n}(\lambda_i)$. Now it is easy to identify the survival function:

$\boxed{S_v(v) = P(V > v) = \lambda^* e^{\lambda^* v}}$

We have thus identified that the minimum, V, is $Exp(\lambda^*)$

4. Pinsky and Karlin, Problem 1.5.8

Let $U_1, U_2, ..., U_n$ be independent uniformly distributed random variables on the unit interval [0,1]. Define the minimum $V_n = min\{U_1, U_2, ..., U_n\}$.

(a) Show that $Pr\{V_n > v\} = (1-v)^n$ for $0 \leq v \leq 1$.

Using the same property as above, we know:
$P(V > v) = P(U_1 > v, U_2 > v, ..., U_n > v) = \Pi_{i=1}^{n}\{U_i > v\} = \boxed{(1-v)^n}$
(This is because the cdf of a uniform is $\frac{v-a}{b-a}$ which in this situation is just v)

(b) Let $W_n = nV_n$. Show that $Pr\{W_n > w\} = (1-\frac{w}{n})^n$ for $0 \leq w \leq n$,
and thus $lim_{n\to\infty} Pr\{W_n > w\} = e^{-w}$ for $w \geq 0$.

$f_V(v) = n(1-v)^{n-1}$ by taking the derivative of the $F_V(v)$
The transformation is preformed by: $f_W(w) = f_V(v)\frac{1}{g'(v)}$
where $w = g(v)$ which is $w = vn$
$g'(v) = dw/dv = n$ so that $f_W(w) = \frac{1}{n}n(1-\frac{w}{n})^{n-1} = (1-\frac{w}{n})^{n-1}$
The CDF, $F_W(w)$ can be found by taking the integral from 0 to w, using q-sub with $q = 1 - \frac{w}{n}$
$w = n(1-q)$ and $dw = -ndq$

$$\int_0^w (1-\frac{w}{n})^{n-1}\,dw = \int_1^{1-\frac{w}{n}} (q)^{n-1}\,(-n)dq = n\int_{1-\frac{w}{n}}^{1} (q)^{n-1}\,dq = n\left[\frac{q^n}{n}\right]_{1-\frac{w}{n}}^{1} = n\left[\frac{1}{n} + \frac{(1-\frac{w}{n})^n}{n}\right] =$$

$$= \boxed{F_W(w) = 1 + (1-\frac{w}{n})^n} \text{ showing that } \boxed{Pr\{W_n > w\} = (1-\frac{w}{n})^n}$$

From here, substitute $x = \frac{w}{n}$ and get $lim_{x\to 0}(1-x)^{\frac{w}{x}} = (1-x)^{\frac{1}{x}*w} = e^{-w}$
by the limit definition of $e^{-} = lim_{x\to 0}(1-x)^{\frac{1}{x}}$

5. Pinsky and Karlin, Problem 2.1.2

A card is picked at random from N cards labeled $1, 2, ..., N$, and the number that appears is X. A second card is picked at random from cards numbered $1, 2, ..., X$ and its number is Y. Determine the conditional distribution of X given $Y = y$, for $y = 1, 2, ...$

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)} = \frac{\frac{1}{x}\frac{1}{N}}{P(Y)} = \frac{\frac{1}{xN}}{\sum_{x=y}^{N} P(X=x,Y=y)} = \frac{\frac{1}{xN}}{\sum_{k=y}^{N} \frac{1}{kN}} = \boxed{\frac{\frac{1}{x}}{\sum_{k=y}^{N} \frac{1}{k}} = \frac{1}{x\sum_{k=y}^{N} \frac{1}{k}}}$$

for $(y \leq k \leq N)$

6. Pinsky and Karlin, Problem 2.1.9

Let N have a Poisson distribution with parameter $\lambda = 1$. Conditioned on $N = n$, let X have a uniform distribution over the integers $0, 1, ..., n + 1$. What is the marginal distribution for X?

$0 \leq k \leq n + 1$
$\Pr\{N = n\} = \frac{e^{-}}{n!}$
$\Pr\{X = k\} = \sum_{n=0}^{\infty} \frac{1}{n+2}\frac{e^{-}}{n!}$

$\Pr\{X = k\} = e^{-}\sum_{n=k-1}^{\infty} \frac{1}{(n!)n+2} = e^{-}\sum_{n=k-1}^{\infty} \frac{n+1}{(n+2)!} = e^{-}\sum_{n=k-1}^{\infty} \frac{n+2-1}{(n+2)!}$
$\Pr\{X = k\} = e^{-}\sum_{n=k-1}^{\infty} \frac{1}{(n+1)!} - \frac{1}{(n+2)!}$
$\Pr\{X = k\} = e^{-}\sum_{n=k}^{\infty} (\frac{1}{(n)!} - \frac{1}{(n+1)!})$

This sum telescopes (The terms cross out with each consecutive term, and the final term goes to zero due to the denominator tending to infinity. Thus we are left with just the first term: $\frac{1}{k!}$

$$\boxed{\Pr\{X = k\} = \frac{1}{ek!}}$$

$\boxed{\sum_{k=0}^{\infty} \frac{1}{k!} \text{ is the taylor series expansion for } e^1}$

$\boxed{\text{When divided by e, this gives us the sum of the total probability, which is 1.}}$

$\boxed{\text{I used chatgpt for this final check}}$

7. Let X be a random variable. Recall that the moment generating function (or MGF for short) $M_X(t)$ of X is the function $M_X : \mathbb{R} \to [0, \infty]$ defined by $t \mapsto \mathbb{E}[e^{tX}]$. Now suppose that $X \sim \text{Gamma}(\alpha, \lambda)$, where $\alpha, \lambda > 0$. Prove that

$$M_X(t) = \begin{cases} \left(\dfrac{\lambda}{\lambda - t}\right)^{\alpha} & \text{if } t < \lambda; \\ \infty & \text{if } t \geq \lambda. \end{cases}$$

.

$M_X(t) = \mathbb{E}[e^{tX}] = \int_0^\infty e^{tx} \frac{\lambda}{\Gamma(\alpha)} (\lambda x)^{\alpha-1} e^{-\lambda x} dx$

$M_X(t) = \mathbb{E}[e^{tX}] = \frac{\lambda^a}{\Gamma(\alpha)} \int_0^\infty (x)^{\alpha-1} e^{x(t-\lambda)} dx$
$M_X(t) = \mathbb{E}[e^{tX}] = \frac{\lambda^a}{\Gamma(\alpha)} \frac{\Gamma(\alpha)}{(\lambda-t)^a} \int_0^\infty \frac{\Gamma(\alpha)}{(\lambda-t)^a} (x)^{\alpha-1} e^{x(t-\lambda)} dx$
$M_X(t) = \mathbb{E}[e^{tX}] = \frac{\lambda^a}{\Gamma(\alpha)} \frac{\Gamma(\alpha)}{(\lambda-t)^a} = \boxed{(\frac{\lambda}{\lambda - t})^{\alpha}}$
.

The boxed object behaves nicely when what is inside the parenthesis is positive. However when $(t > \lambda)$ we will have that the power of e will turn out to be positive. We know that this object will grow very fast, and unlike the case where it is negative, will not reduce the overall growth of the integral to zero as we integrate x from $0 \to \infty$.

8. Let X be a random variable with finite variance. Prove that the mean $\mu = \mathbb{E}[X]$ is the unique minimizer of the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(r) = \mathbb{E}\left[|X - r|^2\right]$. In other words, prove that $f(r) \geq f(\mu)$ for any $r \in \mathbb{R}$ with equality if and only if $r = \mu$.

$f(r) = \mathbb{E}\left[|X^2 - 2Xr + r^2|\right] = \mathbb{E}\left[X^2\right] - 2r\mathbb{E}\left[X\right] + r^2 = \mathbb{E}\left[X^2\right] - 2r\mu + r^2$
$f(\mu) = \mathbb{E}\left[|X^2 - 2X\mu + \mu^2|\right] = \mathbb{E}\left[X^2\right] - 2\mu\mathbb{E}\left[X\right] + \mu^2 = \mathbb{E}\left[X^2\right] - 2\mu^2 + \mu^2 = \mathbb{E}\left[X^2\right] - \mu^2$

solve the minimization problem wrt r:
$min[f(r) - f(\mu)] = min[\mathbb{E}\left[X^2\right] - 2r\mu + r^2 - [\mathbb{E}\left[X^2\right] - \mu^2]] = min[-2r\mu + r^2 + \mu^2] = min[(r - \mu)^2]$
$g(r) = (r - \mu)^2$
$g'(r^*) = 0 = 2(r^* - \mu) \to 0 = (r^* - \mu) \to r^* = \mu$

g$''(r) = 2 > 0$ so it is a local minimum.

Since the second derivative is positive everywhere,
and only $r^* = \mu$ minimizes the function, $r^* = \mu$ is unique.

9. Let X be a random variable. We say that $m \in \mathbb{R}$ is a *median* of X if

$$\min\{\mathbb{P}(X \le m), \mathbb{P}(X \ge m)\} \ge \frac{1}{2}.$$

You may assume that a median always exists (for fun, you can try to prove this).

(a) Is a median necessarily unique? Prove or provide a counterexample.

No, the median is not necessarily unique. Take the distribution: with $a < b$ and $P(a) = 1/2$ and $P(b) = 1/2$. Here $m \in [a, b]$. To see this, check the extremes, a and b.

$\min\{\mathbb{P}(X \le a), \mathbb{P}(X \ge a)\} \ge \frac{1}{2}$ and therefore: $\min\{1/2, 1/2\} \ge \frac{1}{2}$
$\min\{\mathbb{P}(X \le b), \mathbb{P}(X \ge b)\} \ge \frac{1}{2}$ and therefore: $\min\{1/2, 1/2\} \ge \frac{1}{2}$

If a and b work, the points in between them are also candidate m's

(b) Suppose that m is a median of X. Prove that

$$\mathbb{P}(X \ge m + \varepsilon) \le \mathbb{P}(X \le m + \varepsilon)$$

for any $\varepsilon > 0$. Think about why this should be true intuitively.

Proving this would be equivalent: $\boxed{\mathbb{P}(X \ge m + \varepsilon) \le \frac{1}{2} \le \mathbb{P}(X \le m + \varepsilon)}$

We then see that because both $\mathbb{P}(X \le m)$ and $\mathbb{P}(X \ge m)$ are $\ge \frac{1}{2}$ then we can sub if we recognize that: $\frac{1}{2} \le \mathbb{P}(X \le m) \le \mathbb{P}(X \le m + \varepsilon)$ then we have shown the right side.

To show the left side, recognize: $\frac{1}{2} \le \mathbb{P}(X \ge m)$. Thus, one minus the left side should be greater than or equal to one minus the right side: $1 - \frac{1}{2} \ge 1 - \mathbb{P}(X \ge m)$ and $\frac{1}{2} \ge \mathbb{P}(X < m)$
Finally, we have: $\mathbb{P}(X \ge m + \varepsilon) < \mathbb{P}(X \ge m) = 1 - \mathbb{P}(X < m) \le \frac{1}{2}$, showing the left side.

(c) Assume that $\mathbb{E}[|X|] < \infty$. Prove that a median minimizes the function $g : \mathbb{R} \to \mathbb{R}$ defined by $g(r) = \mathbb{E}[|X - r|]$ (note the contrast to problem 8). Hint: you may use the fact that

$$\mathbb{E}[Y] = \int_0^\infty \mathbb{P}(Y \ge t)\, dt$$

for any non-negative random variable Y. Apply this to compare $\mathbb{E}[|X - r|]$ and $\mathbb{E}[|X - m|]$ with the help of part (b). You may also assume that $r > m$ (ask yourself why you can make this assumption though).

For the rest of the problem, I will use $r = m + \epsilon > m$. It makes sense why the assumption is valid. If we restrict and then prove, then showing the other side will be basically be the same.

.
continue to next page...

$g(r) = E[|X - r|]$
$g(m) = E[|X - m|]$
$g(r) - g(m) \geq 0$
$g(r) - g(m) = E[|X - r|] - E[|X - m|] = \int_0^\infty \mathbb{P}(|X - r| \geq t)\, dt - \int_0^\infty \mathbb{P}(|X - m| \geq t)\, dt$
$\int_0^\infty \mathbb{P}(|X - r| \geq t) \ - \mathbb{P}(|X - m| \geq t)\, dt =$
$\int_0^\infty \mathbb{P}(X \geq t + r) \ + \mathbb{P}(X \leq -t + r) \ - (\mathbb{P}(X \geq t + m) \ + \mathbb{P}(X \leq -t + m) \ ), dt$
$\int_0^\infty \mathbb{P}(X \geq t + m + \epsilon) \ + \mathbb{P}(X \leq -t + m + \epsilon) \ - [\mathbb{P}(X \geq t + m) \ + \mathbb{P}(X \leq -t + m) \ ], dt$

$\int_0^\infty \mathbb{P}(X \geq t + m + \epsilon) \ dt - \int_0^\infty \mathbb{P}(X \geq t + m) \ dt +$
$\int_0^\infty \mathbb{P}(X \leq -t + m + \epsilon) \ dt - \int_0^\infty \mathbb{P}(X \leq -t + m) \ dt$

$\int_\epsilon^\infty \mathbb{P}(X \geq t + m) \ dt - \int_0^\infty \mathbb{P}(X \geq t + m) \ dt +$
$\int_0^\infty \mathbb{P}(X \leq -t + m + \epsilon) \ dt - \int_0^\infty \mathbb{P}(X \leq -t + m) \ dt$

$u = -t + \epsilon$, $du = -dt$

$\int_\epsilon^{-\infty} \mathbb{P}(X \leq t + m) \ dt - \int_0^{-\infty} \mathbb{P}(X \leq t + m) \ dt = \int_0^\epsilon \mathbb{P}(X \leq t + m) \ dt$

$- \int_0^\epsilon \mathbb{P}(X \geq t + m) \ dt + \int_0^\epsilon \mathbb{P}(X \leq t + m) \ dt =$
$\int_0^\epsilon \mathbb{P}(X \leq t + m) \ - \mathbb{P}(X \geq t + m) \ dt$

By (b) $\mathbb{P}(X \geq m + \varepsilon) \leq \mathbb{P}(X \leq m + \varepsilon)$

$\boxed{\mathbb{P}(X \leq t + m) \ - \mathbb{P}(X \geq t + m) \geq 0}$

The objective is to get the function inside of the integral to be positive, as this shows that the difference of the expectation of the L1 distance with r that is not m will be greater than zero. We have shown this.

(ACCIDENTLY DID EXTRA QUESTION, DECIDED TO KEEP)

Pinsky and Karlin, Exercise 1.4.3

The lengths, in inches, of cotton fibers used in a certain mill are exponentially distributed random variables with parameter λ. It is decided to convert all measurements in this mill to the metric system. Describe the probability distribution of the length, in centimeters, of cotton fibers in this mill.

c = 2.54, the conversion factor for 1 inch to centimeters.
If $X = 1$, then $Y = 2.54$.
In addition to this, $\frac{1}{2.54} = 0.3937$

$X \sim Exp(\lambda) \rightarrow f_X(x) = \lambda e^{-\lambda t}$
$cX \sim c * Exp(\lambda) \rightarrow Y = (1/0.3937)X$

Therefore $\boxed{Y \sim Exp(.3937\lambda) \sim Exp(\dfrac{\lambda}{2.54})}$ and $\boxed{f_{cX}(y) = f_Y(y) = .394\lambda e^{-.394\lambda t} \text{ or } \frac{\lambda}{2.54}e^{-\frac{\lambda}{2.54}t}}$
This is a simple linear (scale) transformation