1. PK syndrome is a rare condition among University Professors that makes them write gruesome images on the chalkboard when they hear their probability class “improvising” on a problem instead of using the tools and techniques that have been put forth in the homeworks and class lectures. There is a test for PK. If a Professor has PK, then 85% of the time the test shows up as positive. However, 15% of the time, the test will show positive anyway even when the tested Professor does not have PK.

What must be the actual prevalence of PK among University Professors in order for the test to be better than merely flipping a coin to decide? Show your work.

Let \( T \) be the event that a professor tests positive. Let \( A \) be the event that the professor actually has PK syndrome, and \( A^c \) be the complement of \( A \) = event that the professor actually does NOT have PK syndrome. Notice that \( A \) and \( A^c \) form an event family.

The problem tells us that
\[
P \left[ T \mid A \right] = .85 \quad \text{and} \quad P \left[ T \mid A^c \right] = .15
\]
and asks us to determine what values of \( P \left[ A \right] \) make \( P \left[ A \mid T \right] > 1/2 \). Now
\[
P \left[ A \mid T \right] = \frac{P \left[ AT \right]}{P \left[ T \right]} = \frac{P \left[ T \mid A \right] P \left[ A \right]}{P \left[ T \right]} = \frac{.85 P \left[ A \right]}{P \left[ T \right]}
\]
We can express \( P \left[ T \right] \) using Bayes formula:
\[
P \left[ T \right] = P \left[ T \mid A \right] P \left[ A \right] + P \left[ T \mid A^c \right] P \left[ A^c \right] = .85 P \left[ A \right] + .15 \left( 1 - P \left[ A \right] \right) = .15 + .7 P \left[ A \right]
\]
and so
\[
P \left[ A \mid T \right] > 1/2 \iff \frac{P \left[ A \right]}{.15 + .7 P \left[ A \right]} > .5 \iff 1.7 P \left[ A \right] > .15 + .7 P \left[ A \right] \iff P \left[ A \right] > .15
\]
So \( \text{the actual prevalence of PK must be greater than 15\% for the test to be better than flipping a coin!} \)

2. A Digital communication channel has a the probability of \( p_b \) in making a bit error (i.e. with probability \( p_b \) a sent “1” is received as a zero by the receiver or vice-versa. The probability of making an error on any given bit is assumed to be independent of what happened when the previous bit was sent.

We need to send 4-bit words over the channel and have two choices:

- We can send the words unadulterated just as 4 bits. In this case, the data is corrupted if any bit is in error, or
- We can use a coding scheme and code the word into a 7 bit codeword and send 7 bits. Because the code has some error correction capability, in this case the data is corrupted only if two bits are in error.
(a) Assume that an error is modeled by exactly one bit being in error. Under this assumption
the probability the data is corrupted for the non-coded case is \((\frac{1}{4})p_b(1-p_b)^3\) which is
obtained quickly using repeated trial theory. Find an expression for the probability that
the data will be corrupted if we use the coding scheme assuming corruption occurs when
exactly two bits are in error.
The expression for the probability that exactly two bits are in error when sending 7 over
the channel is
\[
\binom{7}{2} p_b^2(1-p_b)^5 = 21p_b^2(1-p_b)^5
\]

(b) From your results in part 2a find for what range of \(p_b\) the coding scheme beneficial.
The coding scheme fails to be beneficial (assuming exactly 1 bit and exactly two bits
constitute corrupted data in the uncoded and coded cases, respectively) when
\[
21p_b^2(1-p_b)^5 > 4p_b(1-p_b)^3 \iff \frac{21}{4} p_b(1-p_b)^2 - 1 > 0 \iff 21p_b^3 - 42p_b^2 + 21p_b - 4 > 0
\]
So does the curve \(21x^3 - 42x^2 + 21x - 4\) ever go positive when \(0 \leq x \leq 1\)? Well, at \(x = 0\)
and \(x = 1\) the curve takes the same value of \(-4\). So if it ever does go positive in \([0,1]\)
we have at least one local maximum in there. That is the derivative must vanish
somewhere on \(0 \leq x \leq 1\). The derivative is
\[
63x^2 - 84x + 21
\]
which has roots at \(\frac{84 \pm \sqrt{84^2 - 4(21)(63)}}{126} = \frac{84 \pm 12}{126} = 1, 1/3\). But evaluating the curve at
\(x = 1/3\) gives the value of \(-8/9\) so we conclude the curve NEVER goes positive on \([0,1]\).

The coding is always beneficial under “exactly 1 bit” & “exactly 2 bit” data corruption models.

(c) Derive the error probabilities for coded and non-coded transmission under the more ac-
accurate assumption that data is corrupted in the non-coded case when at least one bit is
corrupted, and in the coded case when at least two bits are corrupted.
For the uncoded case, we notice that if \(E = \) the event that at least one bit is in error,
then \(E^c = \) complement of \(E = \) event that NO bits are in error. Hence
\[
P[E] = (1 - (1-p_b)^4)
\]
It’s also possible to break \(E\) up into the union of mutually exclusive events \(E_1, E_2, E_3\)
and \(E_4\) — the events of exactly 1, 2, 3, and 4 bits in error, respectively. In this case
\[
= \binom{4}{1} p_b(1-p_b)^3 + \binom{4}{2} p_b^2(1-p_b)^2 + \binom{4}{3} p_b^3(1-p_b) + \binom{4}{4} p_b^4
= 4p_b(1-p_b)^3 + 6p_b^2(1-p_b)^2 + 4p_b^3(1-p_b) + p_b^4
\]
You can check that the expressions in (1) and (2) are equal to each other (with a little
algebraic elbow grease).
For the coded case, if \(F = \) event that at least two bits are in error, then \(F^c = \) event that
less than two bits are in error and
\[
F^c = F_0 \cup F_1 = \text{event no bits are in error} \cup \text{event exactly 1 bit is in error}
\]
Because \(F_0\) and \(F_1\) are mutually exclusive (and, of course, because \(P[F] = 1 - P[F^c]\))
we get
\[
P[F] = 1 - (1-p_b)^7 - \binom{7}{1} p_b(1-p_b)^6 = 1 - (1-p_b)^7 - 7p_b(1-p_b)^6
\]
(d) (Optional Exercise) Using your error expressions from part 2c, find the range of \( p_b \) under which coding is beneficial.

So under the “at least 1 bit” and “at least two bits” model of data corruption, coding is beneficial when

\[
P[E] > P[F] \Leftrightarrow 1 - (1 - p_b)^4 > 1 - (1 - p_b)^7 - 7p_b(1 - p_b)^6
\]
\[
\Leftrightarrow - (1 - p_b)^4 > - (1 - p_b)^7 - 7p_b(1 - p_b)^6
\]
\[
\Leftrightarrow (1 - p_b)^4 < (1 - p_b)^7 + 7p_b(1 - p_b)^6
\]
\[
\Leftrightarrow 1 < (1 - p_b)^3 + 7p_b(1 - p_b)^2
\]
\[
\Leftrightarrow 0 < 6p_b^3 - 11p_b^2 + 4p_b
\]
\[
\Leftrightarrow 0 < 6p_b^2 - 11p_b + 4
\]

So, now the question where of \( 0 \leq x \leq 1 \) is the curve \( 6x^2 - 11x + 4 \) positive? At \( x = 0 \) the value is 4 which is positive. And at \( x = 1 \) the value is \(-1\) which is negative. The roots of the quadratic are at \( x = 1/2 \) and \( x = 4/3 \). So, the curve is evidently positive for \( 0 \leq x < 1/2 \) and then negative for \( 1/2 < x \leq 1 \). So

\[
\text{Coding is beneficial when } p_b < 1/2
\]

3. A discrete random variable \( X \) can take on the values \(-1, 2, 3, \) or \( 5 \). The CDF (cumulative distribution function), \( F_X(x) \), for \( X \) is shown in Figure 1 and it takes on the values

![Figure 1: CDF for Random Variable X](image)

0, 1/10, 3/10, 7/10, and 1.
(a) What is the probability $P[X \leq 3/2]$?
This is simply $F_X(3/2) = \frac{1}{10}$

(b) What is the probability $P[-1/2 < X \leq 7/2]$?
This is $F_X(7/2) - F_X(-1/2) = \frac{7}{10} - \frac{1}{10} = \frac{6}{10}$

(c) Compute the expected value $\mu = E[X]$.
The PDF will have $\delta$’s at $x = -1, 2, 3,$ and $5$. The "size" of the delta is the size of the jump in the CDF, so $1/10$ at $x = -1$, then $2/10$ at $x = 2$, then $4/10$ at $x = 3$ and $3/10$ at $x = 5$. So

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) \, dx$$

$$= \int_{-\infty}^{\infty} x \left( \frac{1}{10} \delta(x+1) + \frac{2}{10} \delta(x-2) + \frac{4}{10} \delta(x-3) + \frac{3}{10} \delta(x-5) \right) \, dx$$

$$= \int_{-\infty}^{\infty} x \frac{1}{10} \delta(x+1) \, dx + \int_{-\infty}^{\infty} \frac{2}{10} x \delta(x-2) \, dx + \int_{-\infty}^{\infty} \frac{4}{10} x \delta(x-3) \, dx + \int_{-\infty}^{\infty} \frac{3}{10} x \delta(x-5) \, dx$$

$$= \frac{1}{10} (-1) + \frac{2}{10} (2) + \frac{4}{10} (3) + \frac{3}{10} (5) = \frac{1}{10} (-1 + 4 + 12 + 15)$$

$$= \frac{30}{10} = 3$$

(d) Compute the variance $\sigma^2 = \text{Var}[X]$.
The variance is

$$\text{Var}[X] = E[X^2] - (E[X])^2$$

$$= \int_{-\infty}^{\infty} x^2 f_X(x) \, dx - (3)^2$$

$$= \int_{-\infty}^{\infty} x^2 \frac{1}{10} \delta(x+1) \, dx + \int_{-\infty}^{\infty} \frac{2}{10} x^2 \delta(x-2) \, dx + \int_{-\infty}^{\infty} \frac{4}{10} x^2 \delta(x-3) \, dx + \int_{-\infty}^{\infty} \frac{3}{10} x^2 \delta(x-5) \, dx - 9$$

$$= \frac{1}{10} (-1)^2 + \frac{2}{10} (2)^2 + \frac{4}{10} (3)^2 + \frac{3}{10} (5)^2 - 9 = \frac{1}{10} (1 + 8 + 36 + 75) - 9$$

$$= \frac{120}{10} - 9 = 12 - 9 = 3$$

4. A continuous random variable $Y$ can take on any value in the interval $[0, 5]$, but only takes values in that interval. It is known that it’s PDF (probability density function), $f_Y(y)$, has the form

$$f_Y(y) = Ay(5 - y)$$

for some positive constant $A$.

(a) Determine the value of $A$. Use that value in the parts below.
$A$ is determined by the requirement that a PDF have total “mass” (i.e. area) equal to one. Now

$$\int_{-\infty}^{\infty} f_X(x) \, dx = \int_{0}^{5} 2y(5 - y) \, dy = 125A \left( \frac{5y^2}{2} - \frac{y^3}{3} \right) \bigg|_{0}^{5} = 125A \left( \frac{125}{2} - \frac{125}{3} \right) = 125A/6$$

So $A = \frac{6}{125}$

(b) Find the expectation $\mu = E[Y]$

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y) \, dy = \int_{0}^{5} y \left( \frac{6}{125} y(5 - y) \right) \, dy = \frac{6}{125} \int_{0}^{5} y^2 (5 - y) \, dy = \frac{5}{2}$$
(c) What is the probability $P[Y > 2]$?

$$P[Y > 2] = \int_{2}^{\infty} f_Y(y) \, dy = \int_{2}^{5} \frac{6}{125} y(5 - y) \, dy = \frac{81}{125}$$

Note the RV is purely continuous (no δ’s in the PDF) so we know $P[Y > 2] = P[Y \geq 2]$.

(d) What is the probability $P[1 < Y \leq 3]$?

$$P[1 < Y \leq 3] = \int_{1}^{3} f_Y(y) \, dy = \int_{1}^{3} \frac{6}{125} y(5 - y) \, dy = \frac{68}{125}$$

Again, $Y$ being a purely continuous RV means $P[1 < Y \leq 3] = P[1 \leq Y \leq 3] = P[1 < Y < 3] = P[1 \leq Y < 3]$.

(e) What is the probability $P[Y > 10]$?

This probability is zero since the PDF is zero outside the range $[0, 5]$. 