Fun Probabilistic Questions

1. There are $n$ urns, and the $k$th urn contains $k - 1$ red balls and $n - k + 1$ white balls. You pick an urn at random and remove two balls at random without replacement. Find the probability that:

   (a) the second ball is white,
   (b) the second ball is white, given that the first ball is white.

2. The Monty Hall Problem. In the old game show *Let’s Make a Deal*, the contestant is presented with three doors and must pick one. Behind one door is a real prize and behind the other two are booby prizes. After the contestant chooses a door the host of the show, Monty Hall, reveals what is behind one of the other two doors, and then gives the contestant the option of switching their choice to the remaining door. Should the contestant switch or not? If they do switch, what is the probability of winning?

   *Important*: When Monty reveals what is behind one of the remaining two doors, he always picks a door containing a booby prize. It wouldn’t be very suspenseful if he revealed the real prize.

   You can read more about the long history of the Monty Hall problem on Wikipedia.

Some Special Distributions

3. The exponential distribution with parameter $\lambda > 0$ has the probability density function

   $$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

   (a) Find the mean and variance of this distribution.
(b) Find the moment generating function of this distribution, i.e. let $X$ be a random variable with the exponential distribution and compute

$$m(t) = E[e^{tX}].$$

(c) For what values of $t \in \mathbb{R}$ is $m(t)$ finite?

(d) Use $m(t)$ to compute the mean and variance of the exponential distribution.

(e) Show that the exponential distribution has the memoryless property, that is show that for $t, s \geq 0$ the relation

$$\mathbb{P}(X > t + s | X > t) = \mathbb{P}(X > s)$$

is true. Suppose that $X$ is the lifespan of a light bulb. Use one sentence to summarize what the memoryless property says about the life of the light bulb.

Remark. It can actually be shown that the exponential distribution is essentially the only probability distribution with the memoryless property. For this reason it plays an important role in the study of continuous-time Markov processes.

4. Let $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$ be independent normal random variables.

(i) Show that $E[e^{itX}] = \exp\{it\mu_1 - t\sigma_1^2/2\}$. Hint: you can just integrate $e^{itx}$ against the density of the normal distribution and pretend that it is a real number throughout, but if you want to be really proper you should use some contour integrals.

(ii) Show that the sum is also normal distributed by using characteristic functions. Is this the characteristic function of the answer you got in the first part?

(iii) Compute the probability density of their sum by using convolution.

(iv) What are the parameters of the sum (i.e. what is the mean and variance of the $X+Y$)?

**Functions as Random Variables**

In the first two questions we work on the probability space $([0, 1], \mathcal{B})$. That is we’re pulling real numbers from the interval $[0, 1]$ out of the hat. Let $f : [0, 1] \to \mathbb{R}$ be a positive function such that

$$\int_0^1 f(x) \, dx = 1.$$  

If we define $\mathbb{P} : \mathcal{B} \to [0, 1]$ by

$$\mathbb{P}(A) = \int_A f(x) \, dx,$$

we have a probability measure.
then \( \mathbb{P} \) is a probability measure on \([0,1], \mathcal{B} \). Sometimes it is abbreviated by \( d\mathbb{P}(x) = f(x) \, dx \).

This type of statement is read to say that \( \mathbb{P} \) has a density with respect to the Lebesgue measure \( dx \).

Random variables on this probability space \([0,1], \mathcal{B}, \mathbb{P} \) are just functions \( X : [0,1] \to \mathbb{R} \). Their expectation is

\[
\mathbb{E}[X] = \int_0^1 X(x) \, d\mathbb{P}(x) = \int_0^1 X(x) f(x) \, dx.
\]

The distribution function of a random variable \( X \), under \( \mathbb{P} \), is defined as

\[
F_X(x) = \mathbb{P}(\omega \in \Omega : X(\omega) \leq x) = \mathbb{P}(X \leq x).
\]

5. Consider the case where the density \( f(x) = 1 \). This the uniform (Lebesgue) measure on \([0,1]\) since all numbers have equal weight. Compute the expectations and the distribution functions of the following random variables:

   (i) \( X(x) = x \),
   (ii) \( Y(x) = x^2 \),
   (iii) \( Z(x) = \begin{cases} 
   1, & 0 \leq x \leq 3/4 \\
   -1, & 3/4 < x \leq 1
   \end{cases} \)
   (iv) \( U(x) = \begin{cases} 
   1, & x \in \mathbb{Q} \\
   0, & x \notin \mathbb{Q}
   \end{cases} \)
   (v) \( V(x) = \sin(2\pi x) \)

6. Now suppose that the density function in the last question is \( f(x) = 3x^2 \). Answer the last problem under this density.

## Convergence in Distribution

Recall that a sequence of distributions functions \( F_n \) converge to the distribution function \( F \) if

\[
F_n(x) \to F(x)
\]

at all points \( x \) where \( F \) is continuous. If a random variable \( X_n \) has \( F_n \) as its distribution function, and the random variable \( X \) has \( F \) as its distribution function, then we can say that \( X_n \) converges in law (or in distribution) to \( X \). This is written as \( X_n \xrightarrow{\text{(d)}} X \).

Note that this type of convergence only depends on the distribution of the random variables, not on the random variables themselves (as we saw in class, two different random variables can have the same distribution function).
7. If \( X \) is a discrete random variable taking values in the integers, meaning that if \( k \geq 0 \) then \( \mathbb{P}(X = k) = p_k \) for some numbers \( p_k \geq 0 \) with \( \sum_k p_k = 1 \), then to check that \( X_n \xrightarrow{(d)} X \) it is enough to check that

\[
\mathbb{P}(X_n = k) \xrightarrow{n \to \infty} \mathbb{P}(X = k)
\]

for every integer \( k \). Suppose that \( X_n \) has the Binomial\((n, \lambda/n)\) distribution for some \( \lambda > 0 \). Show that

\[
X_n \xrightarrow{(d)} \text{Poisson}(\lambda),
\]

the latter being the Poisson distribution with the parameter \( \lambda \).

This gives an important interpretation of Poisson random variables. Recall that Binomial\((n, \lambda/n)\) is the number of heads one gets when tossing a coin \( n \) times and the probability of each head is \( \lambda/n \). This question says that a Poisson random variable is the number of heads you get when you toss many, many coins, but each one has only a tiny probability of showing up heads. The large number of tosses perfectly balances off against the low probability of success.

8. Let \( X_1, X_2, X_3, \ldots \) be independent random variables, each with the same distribution:

\[
\mathbb{P}(X_n = 0) = \mathbb{P}(X_n = 1) = 1/2.
\]

Now define two new sequences of random variables \( Y_1, Y_2, Y_3, \ldots \) and \( Z_1, Z_2, Z_3, \ldots \), by

\[
Y_n = \sum_{k=1}^{n} X_k 2^{-k}, \quad Z_n = \sum_{k=1}^{n} 2X_k 3^{-k}
\]

(i) What are the possible values of \( Y_1 \)? Of \( Y_2 \)? Of \( Y_3 \)? Of \( Y_n \)?

(ii) What is the distribution function of \( Y_1 \)? Of \( Y_2 \)? Of \( Y_3 \)? Draw quick sketches of them.

(iii) Based on the last question, can you figure out what the sequence \( Y_n \) converges in distribution to? Hint: what is the limiting distribution function, and then what kind of random variable has that as its distribution function?

(iv) Answer the same three questions for \( Z_n \) instead of \( Y_n \). Hint: this is related to Problem 12 of the last homework.

9. Recall the Happy Meal collecting problem of the last homework. In that example there were \( K \) toys to collect. Let \( X_K \) be the number of days it takes to collect all \( K \) toys. Clearly \( X_K \) is a random variable, and in the last homework you showed that

\[
\mathbb{E}[X_K] = K \sum_{j=1}^{K} \frac{1}{j} \approx K \log K
\]
when $K$ is large. You did this by writing

$$X_K = \sum_{j=0}^{K-1} D_j,$$

where $D_j$ is the number of days it takes to obtain the $(j+1)$st new toy after obtaining the $j$th one. The $D_j$ are all independent, and $D_j$ has the geometric distribution with probability of success equal to $(K - j)/K$.

(i) Compute the characteristic function of $X_K$.

(ii) Use the characteristic function to show that

$$\frac{X_K - K \log K}{K} \xrightarrow{(d)} T,$$

as $K \to \infty$, where $T$ is a random variable with distribution function

$$\mathbb{P}(T \leq x) = \exp(-e^{-x}), \quad -\infty < x < \infty.$$

10. Let $\epsilon_n$ be an iid sequence of random variables with $\mathbb{P}(\epsilon_n = 0) = \mathbb{P}(\epsilon_n = 1) = 1/2$. Let $Y_n$ be an iid sequence of exponential random variables with parameter $\lambda > 0$.

(i) Compute the characteristic function of $\epsilon_n$.

(ii) Compute the characteristic function of $Y_n$.

(iii) Compute the characteristic function of $\epsilon_nY_n$.

(iv) Define a new sequence of random variables $X_n$ by the relationship

$$X_{n+1} = \frac{1}{2}X_n + \epsilon_nY_n.$$ 

This is a recursion equation for the $X_n$, the value at time $n + 1$ is determined by the value at time $n$ plus some independent randomness (the $\epsilon_nY_n$ term). To understand what is going on first ignore the $\epsilon_nY_n$ term. Then the equation $X_{n+1} = \frac{1}{2}X_n$ says that at each time the value of $X$ is cut in half. What the second term $\epsilon_nY_n$ does is to add the independent quantity $Y_n$ onto $X$ after the halving operation, but the $Y_n$ is only added part of the time (when $\epsilon_n = 1$).

Compute the characteristic function of $X_{n+1}$ in terms of the characteristic function of $X_n$ and the answer from the last part. Use this to show that the sequence $X_n$ converges in distribution as $n \to \infty$.

11. Let $X_1, X_2, X_3, \ldots$ be iid random variables with common distribution function $F$. Let $M_n = \max_{m \leq n} X_m$. Show that

(i) $M_n$ has distribution function $\mathbb{P}(M_n \leq x) = F(x)^n$, 

(ii) Compute the characteristic function of $M_n$. 

(iii) Show that $M_n \xrightarrow{(d)} M$, where $M$ is a random variable with distribution function $F$.
(ii) if \( F(x) = 1 - x^{-\alpha} \) for \( x \geq 1 \), where \( \alpha > 0 \), then for \( y > 0 \) we have
\[
\mathbb{P}(M_n/n^{1/\alpha} \leq y) \to \exp(-y^{-\alpha}) \quad \text{as } n \to \infty,
\]

(iii) if \( F(x) = 1 - |x|^\beta \) for \(-1 \leq x \leq 0\) where \( \beta > 0 \), then for \( y < 0 \)
\[
\mathbb{P}(n^{1/\beta}M_n \leq y) \to \exp(-|y|^\beta) \quad \text{as } n \to \infty,
\]

(iv) if \( F(x) = 1 - e^{-x} \) for \( x \geq 0 \), then for all \( y \in (-\infty, \infty) \),
\[
\mathbb{P}(M_n - \log n \leq y) \to \exp(-e^{-y}) \quad \text{as } n \to \infty.
\]

The limits that appear on the right hand side of (ii), (iii), and (iv) are called **extreme value distributions**. They are the limiting distribution functions of the maxima of independent random variables. Note that in the context of this question they appeared

Integration

12. Prove **Chebyshev’s inequality** which says that if \( X \) is a random variable with \( \mathbb{E}[|X|] < \infty \) then
\[
\mathbb{P}(|X| \geq a) \leq \frac{1}{a} \mathbb{E}[|X|].
\]

Hint: use that if \( Y \) and \( Z \) are random variables with \( Y \leq Z \) then \( \mathbb{E}[Y] \leq \mathbb{E}[Z] \), and that \( \mathbb{P}(|X| \geq a) = \mathbb{E}[\mathbf{1}\{|X| \geq a\}] \). Can you think of a random variable that bounds \( \mathbf{1}\{|X| \geq a\} \)?

13. Recall that the sums
\[
\sum_{n=1}^{\infty} \frac{1}{n^p}
\]
converge for \( p > 1 \) and diverge for \( p \leq 1 \). We consider exclusively the \( p > 1 \) case. Let \( C_p = \sum_{n>1} n^{-p}, \) and let \( Y \) be a discrete random variable taking values in \( \{1, 2, 3, \ldots\} \), with probabilities given by
\[
\mathbb{P}(Y = n) = \frac{n^{-p}}{C_p}.
\]

(a) Show that \( Y \) has finite mean only if \( p > 2 \).

(b) Show that \( Y \) has finite variance only if \( p > 3 \).

(c) For \( r > 0 \), we say that \( Y \) has an \( r \)th moment if \( \mathbb{E}[Y^r] < \infty \). For a given value of \( p \) what is the largest moment that \( Y \) has?