ECE 361 Probability for Engineers (Fall, 2016)
Homework Solutions 5

Please answer the following questions:

1. (5 points) Let X be a continuous RV with support $\mathcal{X} = [0,1]$ and PDF $f_X(x) = cx(1-x)$ for $x \in \mathcal{X}$ and some constant c.

   - Give the constant c such that f_X is a valid PDF.
     Solution. $c = 6$ since $\int_{\mathcal{X}} f_X(x)\mathrm{d}x = \int_0^1 6x(1-x)\mathrm{d}x = 1$.
   - Give the CDF $F_X(x)$.
     Solution. $F_X(x) = (3-2x)x^2$ for $x \in [0,1]$, since $\int_{-\infty}^x f_X(t)\mathrm{d}t = \int_0^x 6t(1-t)\mathrm{d}t = (3-2x)x^2$. Of course $F_X(x) = 0$ for $x < 0$ and $F_X(x) = 1$ for $x > 1$.
   - Find $\mathbb{E}[X]$.
     Solution. By symmetry of $f_X(x)$ we expect $\mathbb{E}[X] = \frac{1}{2}$. Formally, $\int_{\mathcal{X}} x f_X(x)\mathrm{d}x = \int_0^1 6x^2(1-x)\mathrm{d}x = \frac{1}{2}$.
   - Find $\mathrm{Var}(X)$.
     Solution. We first compute $\mathbb{E}[X^2] = \int_{\mathcal{X}} x^2 f_X(x)\mathrm{d}x = \int_0^1 6x^3(1-x)\mathrm{d}x = \frac{3}{10}$, and then compute $\mathrm{Var}(X) = 3/10 - (1/2)^2 = 1/20$.
   - Plot $f_X(x)$ vs. x and $F_X(x)$ vs. x over $x \in \mathcal{X}$ on the same plot. Indicate $\mathbb{E}[X]$ on the x-axis.
     Solution.

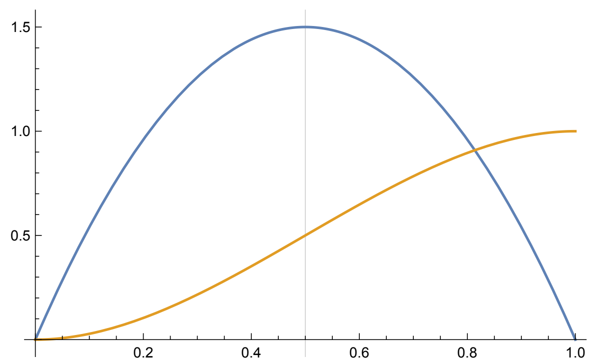

Figure 1: Problem 1.

2. (5 points) Let X be a continuous RV with support $\mathcal{X} = (0,\infty)$ and PDF $f_X(x) = cx^2\mathrm{e}^{-x}$ for $x \in \mathcal{X}$ and some constant c.

   - Give the constant c such that f_X is a valid PDF.
     Solution. $c = \frac{1}{2}$ since $\int_{\mathcal{X}} f_X(x)\mathrm{d}x = \int_0^\infty \frac{1}{2}x^2\mathrm{e}^{-x}\mathrm{d}x = 1$.
   - Give the CDF $F_X(x)$.
     Solution. $F_X(x) = 1 - \frac{1}{2}\mathrm{e}^{-x}(2 + x(2+x))$ for $x \geq 0$, since $\int_{-\infty}^x f_X(t)\mathrm{d}t = \int_0^x \frac{1}{2}t^2\mathrm{e}^{-t}\mathrm{d}t = 1 - \frac{1}{2}\mathrm{e}^{-x}(2 + x(2+x))$. Of course $F_X(x) = 0$ for $x < 0$.
   - Find $\mathbb{E}[X]$.
     Solution. $\int_{\mathcal{X}} x f_X(x)\mathrm{d}x = \int_0^\infty \frac{1}{2}x^3\mathrm{e}^{-x}\mathrm{d}x = 3$.

- Find Var(X).
  Solution. We first compute $\mathbb{E}[X^2] = \int_{\mathcal{X}} x^2 f_X(x)\mathrm{d}x = \int_0^\infty \frac{1}{2}x^4 \mathrm{e}^{-x}\mathrm{d}x = 12$, and then compute $\mathrm{Var}(X) = 12 - 3^2 = 3$.
- Plot $f_X(x)$ vs. x and $F_X(x)$ vs. x over $x \in (0, 10]$ on the same plot. Indicate $\mathbb{E}[X]$ on the x-axis.
  Solution.

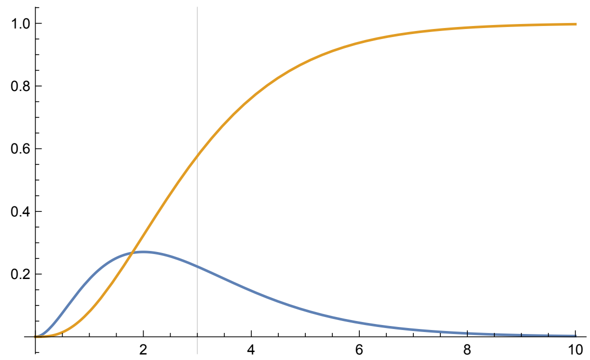

Figure 2: Problem 2.

3. (5 points) Let X be a continuous RV with support $\mathcal{X} = (-\infty, +\infty)$ and PDF $f_X(x) = c\left(1 + x^2\right)^{-1}$ for $x \in \mathcal{X}$ and some constant c.

- Give the constant c such that f_X is a valid PDF.
  Solution. $c = \frac{1}{\pi}$ since $\int_{\mathcal{X}} f_X(x)\mathrm{d}x = \int_{-\infty}^{\infty} \frac{1}{\pi(1+x^2)}\mathrm{d}x = 1$.
- Give the CDF $F_X(x)$.
  Solution. $F_X(x) = \frac{1}{2} + \frac{1}{\pi}\arctan(x)$ for $x \in \mathbb{R}$, since $\int_{-\infty}^{x} f_X(t)\mathrm{d}t = \int_{-\infty}^{x} \frac{1}{\pi(1+t^2)}\mathrm{d}t = \frac{1}{2} + \frac{1}{\pi}\arctan(x)$.
- Plot $f_X(x)$ vs. x and $F_X(x)$ vs. x over $x \in [-5, +5]$ on the same plot.
  Solution.

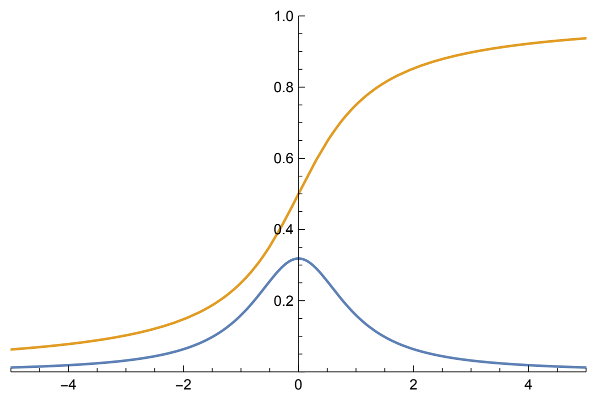

Figure 3: Problem 3.

4. (5 points) The CDF F_X of a continuous RV X may be used to *generate* random variates from that distribution through a technique called *CDF inversion*. The basic idea is to solve the equation $u = F_X(x)$ for x, say $x = F_X^{-1}(u)$, and to then generate $F_X^{-1}(U)$, for $U \sim \mathrm{Uni}([0,1])$ a continuous uniform RV. The justification for this process is the fact that the event $\{X \leq x\}$ is equivalent to the event $\{g(X) \leq g(x)\}$ for any function

$g : \mathcal{X} \to \mathbb{R}$ that is *increasing*, i.e., if $x < y$ then $g(x) < g(y)$. Recall *all* CDFs are *nondecreasing*, i.e., $x < y$ implies $F_X(x) \leq F_X(y)$ (note the nonstrict inequality), but some are in fact (strictly) *increasing*. If $F_X(x)$ is increasing then observe, with $Y = F_X^{-1}(U)$:

$$F_Y(x) = \mathbb{P}(Y \leq x) = \mathbb{P}(F_X^{-1}(U) \leq x) = \mathbb{P}(F_X(F_X^{-1}(U)) \leq F_X(x)) = \mathbb{P}(U \leq F_X(x)) = F_X(x). \quad (1)$$

Here, we have used two key facts. First, if $F_X(x)$ is increasing with inverse $F_X^{-1}(u)$ then $F_X(F_X^{-1}(u)) = u$ (this is the defining property of a function and its inverse). Second, if $U \sim \mathrm{Uni}([0,1])$ then $F_U(u) = \mathbb{P}(U \leq u) = u$ for $u \in [0,1]$. The above proof shows that the CDF for the RV $Y = F_X^{-1}(U)$ equals the CDF $F_X(x)$, i.e., the procedure of generating a *uniform* RV $U \sim \mathrm{Uni}([0,1])$ and computing $Y = F_X^{-1}(U)$ yields a RV with distribution $F_X(x)$, as desired.

We will verify this claim by constructing in this exercise a large collection of independent and identically distributed samples $(X_1, \ldots, X_N)$, with each $X_i \sim F_X$ for $i \in [N]$, and computing the *empirical distribution* for this sample. This exercise is our first foray into *statistics*. First, sort the N samples such that $X_1 < X_2 \ldots < X_N$; note it is safe to assume a strict inequality since the CDF is for a continuous RV and there is zero probability of two samples taking the same value. Second, assign to point i the value i/N for $i \in [N]$. Now connect the points with the continuous function $\tilde{F}_X(x)$ which takes value:

$$\tilde{F}_X(x) = \sum_{i=1}^{N+1} \frac{i-1}{N} \mathbf{1}(x_{i-1} < x \leq x_i). \quad (2)$$

where $x_0 = -\infty$ and $x_{N+1} = +\infty$. Here, the notation $\mathbf{1}(A)$ equals 1 if A is true and 0 if A is false, and is called an *indicator function*. Thus $\tilde{F}_X(x)$ is a sum of indicators and precisely one of these N indicators is true for each $x \in \mathbb{R}$.

Consider the *exponential* distribution $F_X(x) = 1 - \mathrm{e}^{-\lambda x}$ for $x \in \mathcal{X} = [0, +\infty)$, with parameter $\lambda > 0$. Observe $F_X(x)$ is strictly increasing and therefore we can use the CDF inversion technique to generate exponential variates. Please do the following:

- Find $F_X^{-1}(u)$ for $F_X(x) = 1 - \mathrm{e}^{-\lambda x}$ for arbitrary $\lambda > 0$.
  Solution. Solving $F_X(x) = u$ for x gives $1 - \mathrm{e}^{-\lambda x} = u$, or $x = -\log(1-u)/\lambda$.
- Fix $\lambda = 1$ and generate $N = 25$ independent and identically distributed variates $X_1, \ldots, X_N$ with $X_i \sim F_X^{-1}(U)$ for $U \sim \mathrm{Uni}([0,1])$ using your computer's random number generator. E.g., Matlab's `rand` or Mathematica's `Random` or Python's `random`.

  Solution. Using Mathematica I achieve this goal, with sorted values, via the following command:

  ```
  X4[l_,Nn_]:=AssociationThread[Sort[-Log[1-RandomReal[{0,1},Nn]]/l,Range[Nn]/Nn];
  ```
- Sort these values so $X_1 < \cdots < X_N$ and plot the empirical distribution $\tilde{F}_X$ along with the given distribution $F_X(x) = 1 - \mathrm{e}^{-\lambda x}$ for $\lambda = 1$ over $x \in [0, 6]$.
  Solution.
- Repeat the above steps for $\lambda = 2$, plotting over $x \in [0, 3]$.
  Solution.

5. (5 points) Consider a pair of independent and identically distributed uniform RVs (X, Y) with $X \sim \mathrm{Uni}([0,1])$ and $Y \sim \mathrm{Uni}([0,1])$. Think of each random pair (X, Y) as a random point on the unit square $[0,1] \times [0,1]$ on the (x, y)-plane. Let the RV $Z = \sqrt{X^2 + Y^2}$ be the random distance of that random point from the origin. Please do the following:

   - Find the support $\mathcal{Z}$ for the RV Z.
     Solution. $\mathcal{Z} = [0, \sqrt{2}]$.

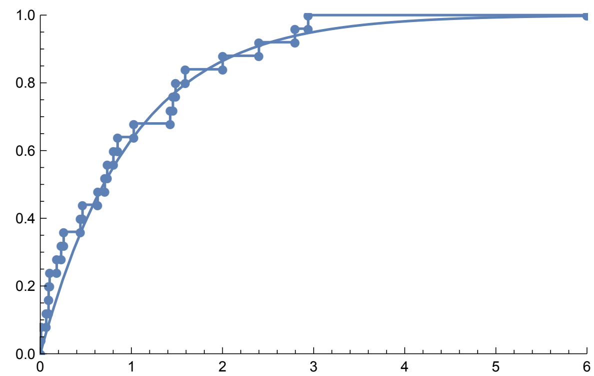

Figure 4: Problem 4 with $\lambda = 1$.

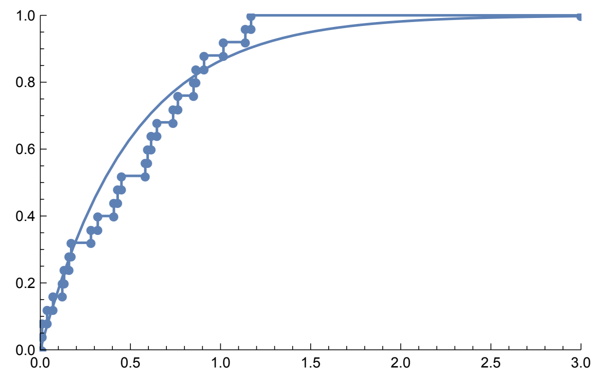

Figure 5: Problem 4 with $\lambda = 2$.

- Find the CDF $F_Z(z)$ for the RV Z. *Hint: divide the support into two sub-intervals and consider the geometry of the problem. The event $\{Z \le z\}$ for z small corresponds to the pair (X,Y) being found in the quarter-circle centered at the origin of radius z. Consider the geometry for z "large" as well.*

  Solution. Divide the support into two sub-intervals: $[0,1)$ and $[1,\sqrt{2}]$. First, for $z \in [0,1)$:

$$\begin{aligned} F_Z(z) &= \mathbb{P}(Z \le z) \\ &= \mathbb{P}(\sqrt{X^2+Y^2} \le z) \\ &= \frac{\pi}{4} z^2 \end{aligned} \tag{3}$$

  Here, we have recognized the event $\{Z \le z\}$ is equivalent to the event $\{(X,Y) \in b(o,z)\}$, where $b(o,z) = \{(x,y) \in \mathbb{R}^2 : \sqrt{x^2+y^2} \le z\}$ is the disk of radius z centered at the origin. The intersection of this disk with the unit box $[0,1]^2$ is a quarter circle with area $\pi z^2/4$. This is shown in the left panel of the figure.

  Next, consider $z \in [1,\sqrt{2}]$, and consider the event $\{Z > z\}$. As shown in the right panel of the figure, this corresponds to the event that (X,Y) is in the region "above" the circle of radius z but in the box $[0,1]^2$. The area of this region is the area of the box with side length $1-\sqrt{z^2-1}$ minus the area of the quarter circle of the same radius:

$$\mathbb{P}(Z > z) = (1-\sqrt{z^2-1})^2 - \frac{\pi}{4}(1-\sqrt{z^2-1})^2 = (1-\sqrt{z^2-1})^2 \left(1 - \frac{\pi}{4}\right). \tag{4}$$

Combining:

$$F_Z(z) = \begin{cases} \frac{\pi}{4} z^2, & 0 \leq z \leq 1 \\ 1 - \left(1 - \frac{\pi}{4}\right)(1 - \sqrt{z^2-1})^2, & 1 \leq z \leq \sqrt{2} \end{cases} \tag{5}$$

- Find the PDF $f_Z(z)$ for the RV Z.
  Solution. Differentiating gives

$$f_Z(z) = \begin{cases} \frac{\pi}{2} z, & 0 \leq z \leq 1 \\ 2\left(1 - \frac{\pi}{4}\right) z \left(\frac{1}{\sqrt{z^2-1}} - 1\right), & 1 \leq z \leq \sqrt{2} \end{cases} \tag{6}$$

- Find $\mathbb{E}[Z]$.
  Solution. Integration yields

$$\int_0^{\sqrt{2}} z f_Z(z) \mathrm{d}z = \frac{1}{12}(8 - (4-\pi)\sqrt{2} + (12 - 3\pi)\mathrm{arcsinh}(1)) \approx 0.754647. \tag{7}$$

where arcsinh is the inverse hyperbolic sine function, with $\mathrm{arcsinh}(1) \approx 0.881374$.

- Plot the CDF $F_Z(z)$ and the PDF $f_Z(z)$ vs. $z \in \mathcal{Z}$. Label the point $\mathbb{E}[Z]$ on the z-axis.
  Solution.

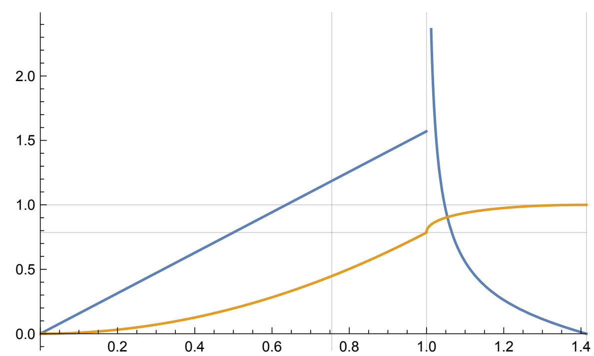

Figure 6: Problem 5. Note the discontinuity in the PDF at $z = 1$.

6. (5 points) Consider a pair of independent and identically distributed uniform RVs (X, Y) with $X \sim \mathrm{Uni}([-1,+1])$ and $Y \sim \mathrm{Uni}([-1,+1])$. Think of each random pair (X, Y) as a random point on the unit square $[-1,+1] \times [-1,+1]$ on the (x,y)-plane. Let the RV $W = \mathbf{1}(\sqrt{X^2+Y^2} \leq 1)$ be the indicator that the random point is at a distance of one or less from the origin, i.e., $W = 1$ if $\sqrt{X^2+Y^2} \leq 1$ and $W = 0$ else. Please do the following:

- Prove that if $\mathbf{1}(A)$ is the function that equals 1 if A is true and 0 if A is false then $\mathbb{E}[\mathbf{1}(A)] = \mathbb{P}(A)$ for any event A. In words, the expected value of an indicator RV for an event A is the probability of A.
  Solution. $\mathbf{1}(A)$ is a Bernoulli RV with $p = \mathbb{P}(A)$, and as such $\mathbb{E}[\mathbf{1}(A)] = \mathbb{P}(A)$.
- Find $\mathbb{E}[W]$.
  Solution. The event $\{W = 1\}$ in this problem corresponds to the event $\{Z \leq 1\}$ in the previous problem, and as such $\mathbb{P}(W = 1) = \mathbb{P}(Z \leq 1) = F_Z(1) = \pi/4$.
- Generate a sequence of $N = 1000$ random (X, Y) values $((X_i, Y_i), i \in [N])$ with each (X_i, Y_i) as above, and compute W_i for each one, as above, yielding a random binary string $(W_i, i \in [N])$. For each $n \in [N]$ define the sequence $(f(n), n \in [N])$, with $f(n) = \frac{1}{n}\sum_{i=1}^n W_i$ the fraction of ones in the first n positions. Plot $f(n)$ vs. $n \in [N]$.
  Solution.

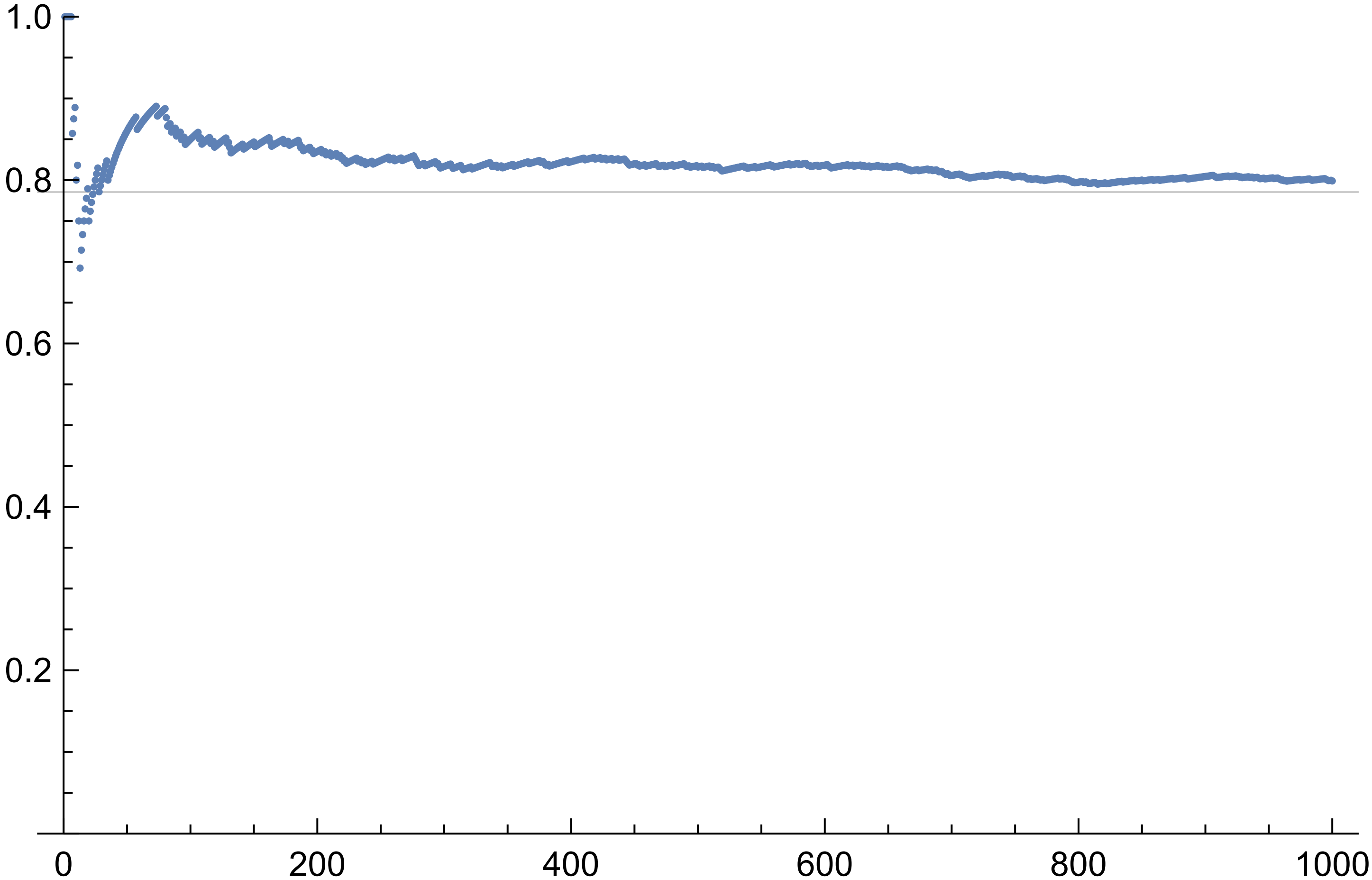

Figure 7: Problem 6. The horizontal gridline is $\pi/4 \approx 0.785398$.

- Guess from your plot the true value of $\lim_{n\to\infty} f(n)$.
  Solution. $f(n) \to \pi/4 \approx 0.785398$.

1. Mathematica code for Problem 1.

```
(* Problem 1 *)
Clear[f1, F1, ex11, ex12, vx1, p1];
f1[x_] := Piecewise[{{6 x (1 - x), 0 <= x <= 1}, {0, x < 0}, {0, x > 1}}];
F1[x_] := Piecewise[{{(3 - 2 x) x^2, 0 <= x <= 1}, {0, x < 0}, {1, x > 1}}];
Integrate[f1[x], {x, 0, 1}]
Assuming[0 < x < 1, FullSimplify[Integrate[f1[t], {t, 0, x}]]]
ex11 = Integrate [x f1[x], {x, 0, 1}]
ex12 = Integrate [x^2 f1[x], {x, 0, 1}]
vx1 = ex12 - (ex11)^2
p1 = Plot[{f1[x], F1[x]}, {x, 0, 1}, GridLines -> {{ex11}, {}}]
Export[NotebookDirectory[] <> "p1.pdf", p1]
```

2. Mathematica code for Problem 2.

```
(* Problem 2 *)
Clear[e, f2, F2, ex21, ex22, vx2, p2];
e := Exp[1];
f2[x_] := Piecewise[{{x^2 Exp[-x]/2, 0 <= x}, {0, x < 0}}];
F2[x_] := Piecewise[{{1 - 1/2 E^-x (2 + x (2 + x)), 0 <= x}, {0, x < 0}}];
Integrate[f2[x], {x, 0, Infinity}]
Assuming[0 < x < 1, FullSimplify[Integrate[f2[t], {t, 0, x}]]]
ex21 = Integrate [x f2[x], {x, 0, Infinity}]
ex22 = Integrate [x^2 f2[x], {x, 0, Infinity}]
vx2 = ex22 - (ex21)^2
p2 = Plot[{f2[x], F2[x]}, {x, 0, 10}, GridLines -> {{ex21}, {}}]
Export[NotebookDirectory[] <> "p2.pdf", p2]
```

3. Mathematica code for Problem 3.

```
(* Problem 3 *)
Clear[f3, F3, ex31, ex32, vx3, p3];
f3[x_] := 1/(\[Pi] (1 + x^2));
F3[x_] := 1/2 + ArcTan[x]/\[Pi];
Integrate[f3[x], {x, -Infinity, Infinity}]
FullSimplify[Integrate[f3[t], {t, -Infinity, x}]]
p3 = Plot[{f3[x], F3[x]}, {x, -5, 5}, PlotRange -> {{-5, +5}, {0, 1}}]
Export[NotebookDirectory[] <> "p3.pdf", p3]
```

4. Mathematica code for Problem 4.

```
(* Problem 4 *)
Clear[F4, F4i, X4, X4a, Nn4, l41, l42, x41, x42, l41, l42, p41, p42];
F4[x_,l_] := Piecewise[{{1 - Exp[-l x], x >= 0}, {0, x < 0}}];
F4i[u_,l_] := Piecewise[{{-Log[1-u]/l, 0<=u<=1}, {Undefined, u<0}, {Undefined, u>1}}];
X4[l_, Nn_] := AssociationThread[Sort[-Log[1 - RandomReal[{0, 1}, Nn]]/l], Range[Nn]/Nn];
X4a[x_,xmax_]:=Module[{x4=Prepend[x,{0->0}]},If[Max[Keys[x4]]<xmax,Append[x4,{xmax->1}],x4]];
```

```
Nn4 = 25;
l41 = 1; l42 = 2;
xmax41 = 6; xmax42 = 3;
x41 = X4a[X4[l41, Nn4], xmax41];
x42 = X4a[X4[l42, Nn4], xmax42];
l41 = ListPlot[x41, InterpolationOrder -> 0, Joined -> True, PlotMarkers -> Automatic];
l42 = ListPlot[x42, InterpolationOrder -> 0, Joined -> True, PlotMarkers -> Automatic];
p41 = Show[Plot[F4[x, l41], {x, 0, xmax41}, PlotRange -> {{0, xmax41}, {0, 1}}], l41]
p42 = Show[Plot[F4[x, l42], {x, 0, xmax42}, PlotRange -> {{0, xmax42}, {0, 1}}], l42]
Export[NotebookDirectory[] <> "p41.pdf", p41]
Export[NotebookDirectory[] <> "p42.pdf", p42]
```

5. Mathematica code for Problem 5.

```
(* Problem 5 *)
Clear[F5, f5, p5];
F5[z_]:=Piecewise[{
{0,z<0},
{pi z^2/4, 0<=z<=1},
{1-(1-Sqrt[z^2-1])^2 (1-pi/4), 1<=z<=Sqrt[2]},
{1, z>1}
}];
f5[z_]:=Piecewise[{
{0, z < 0},
{pi z/2, 0 <= z <= 1},
{2 (1 - pi/4) z (1/Sqrt[z^2 - 1] - 1), 1 <= z <= Sqrt[2]},
{0, z > Sqrt[2]}
}];
ex51 = Integrate[z f5[z], {z, 0, Sqrt[2]}]
p5=Plot[{f5[z],F5[z]},{z,0,Sqrt[2]},
PlotRange->{{0,Sqrt[2]},Automatic},
GridLines->{{1,ex51,Sqrt[2]},{pi/4,1}}
]
Export[NotebookDirectory[] <> "p5.pdf", p5]
```

6. Mathematica code for Problem 6.

```
(* Problem 6 *)
Clear[XY6, W6, N6, p6];
XY6[Nn_] := RandomVariate[UniformDistribution[], {Nn, 2}];
W6[xy_] := Map[If[Norm[#] <= 1, 1, 0] &, xy];
N6 = 1000;
p6 = ListPlot[Accumulate[W6[XY6[N6]]]/Range[N6],
GridLines -> {{}, {pi/4}},
PlotRange -> {Automatic, {0, 1}}
]
Export[NotebookDirectory[] <> "p6.pdf", p6]
```