Problem 1. Let $N$ be a Poisson random variable with mean $\lambda$. Determine the Chernoff bound for $N$. Compare the Chernoff bound for $\lambda = 10$ to the actual value of $P\{N > 20\}$, which you can numerically evaluate using Matlab.

Solution: The Chernoff bound for any nonnegative random variable $N$ is given by the right-hand side of

$$P\{N \geq c\} \leq \inf_{\theta \geq 0} e^{-\theta c} E[e^{\theta N}].$$

This is obtained by first calculating $e^{-\theta c} E[e^{\theta N}]$, differentiating it, and then setting the derivative to zero to obtain the minimum value.

For a Poisson random variable,

$$E[e^{\theta N}] = \sum_{k=0}^{\infty} e^{\theta k} P\{N = k\} = \sum_{k=0}^{\infty} e^{\theta k} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} e^{\theta k} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda e^\theta} = e^{\lambda (e^\theta - 1)}.$$

We can use calculus to minimize $e^{-\theta c} e^{\lambda (e^\theta - 1)}$ over $\theta \geq 0$; this yields $\theta = \log(\frac{c}{\lambda})$, which yields the Chernoff bound

$$P\{N \geq c\} \leq \left(\frac{\lambda}{c}\right)^c e^{-\lambda}.$$

For $c = 20$, the Chernoff bound is 0.0210 while exact the calculation from Matlab gives 0.0072. The Markov and Chebyshev bounds are 2 (very loose!) and 0.275, respectively.

Problem 2. Consider the measurable space $(\Omega, \mathcal{F})$, and let $\mathcal{D}$ be a sub $\sigma$-algebra. Let $X$ and $Y$ be $\mathcal{D}$-measurable random variables.

a) Show that $aX$ is a $\mathcal{D}$-measurable random variable, where $a$ is a real scalar.

b) Show that $X + Y$ is a $\mathcal{D}$-measurable random variable.

Solution:
a) We will show that for any \( a \in \mathbb{R} \), \((aX)^{-1}((\infty, r)) \in \mathcal{D}\), for all \( r \in \mathbb{R} \).

Note, however, that \((aX)^{-1}((\infty, r)) \cap (\infty, r/a)\). To explicitly see this, note that if \( \omega \in (aX)^{-1}((\infty, r)) \) then \( (aX)(\omega) \in (\infty, r) \) or \( X(\omega) \in (\infty, r/a) \). Thus, \( \omega \in X^{-1}((\infty, r/a)) \).

On the other hand, if \( \omega \in X^{-1}((\infty, r/a)) \), then \( X(\omega) \in (\infty, r/a) \) or equivalently \( aX(\omega) \in (\infty, r) \); thus, \( \omega \in (aX)^{-1}((\infty, r)) \). Hence, \((aX)^{-1}((\infty, r)) = X^{-1}((\infty, r/a)) \in \mathcal{D}\) since \( X \) is \( \mathcal{D} \)-measurable.

b) We will show that \((X + Y)^{-1}((\infty, r)) \in \mathcal{D}\) for all \( r \in \mathbb{R} \). For simplicity, we equivalently write \( \{X + Y < r\} \) for \((X + Y)^{-1}((\infty, r)) \). Now, \( \{X + Y < r\} = \bigcup_{i=1}^{\infty} \{X < q_i\} \cap \{Y < r - q_i\} \), where \( q_1, q_2, \ldots \) is an enumeration of the rational numbers. We can do this because the set of rational numbers is a countable set and the rational numbers are dense in \( \mathbb{R} \). But for each \( i \), \( \{X < q_i\} \cap \{Y < r - q_i\} \in \mathcal{D} \), since \( X \) and \( Y \) are each \( \mathcal{D} \)-measurable. Thus, \( \bigcup_{i=1}^{\infty} \{X < q_i\} \cap \{Y < r - q_i\} \in \mathcal{D} \). Thus, \( X + Y \) is \( \mathcal{D} \)-measurable.

Problem 3. Use the built-in Matlab random number generator (the Matlab command is \texttt{rand}) to generate samples of an exponentially-distributed random variable \( Z \). Namely, we want to generate random numbers that are mutually independent and each of them is distributed according to the probability density function \( f_Z(z) = \mu^{-1} \exp(-z/\mu)u(z) \). Note that you can use \texttt{rand(k,l)} to generate a \( k \times l \) array of independent \([0,1]\)-valued uniformly distributed r.v.’s at once.

1. Show that if \( X \) is a uniformly-distributed r.v. in \([0,1] \), then the new r.v. \( Z \) defined by

\[
Z = -\mu \log(1 - X)
\]

is an exponentially-distributed r.v. with parameter \( \mu \).

2. Assuming that \( \mu = 1 \), use the result in (a) to generate \( n = 1000 \) samples of \( f_Z \). Let \( \mathbf{Z} = [Z_1, Z_2, \ldots, Z_n] \) denote the array of identically-distributed and mutually independent samples.

3. Compute \( \max_{1 \leq i \leq n}(Z_i) \), \( \min_{1 \leq i \leq n}(Z_i) \), and the arithmetic mean \( \bar{\mu} \) of \( Z_1, \ldots, Z_n \). (Use the Matlab functions \texttt{max}, \texttt{min} and \texttt{mean}).

4. Use the histogram generation capability of Matlab to generate an empirical estimate of the density function \( f_Z \) using the above 1000 samples. To achieve this task, follow the procedure outlined below.

   (a) Generate the bin array \( \mathbf{B} = [0, B_1, \ldots, B_m] \), where \( m = 100 \) is the total number of bins.

     Use the uniform spacing

\[
B_{i+1} - B_i = \delta,
\]

where

\[
\delta = \max_{1 \leq i \leq n}(Z_i)/m.
\]

(You may use the syntax \( \mathbf{B} = [0: \delta : \max(\mathbf{Z})] \).)

(b) Perform \( \mathbf{H} = \text{hist}(\mathbf{Z}, \mathbf{B}) \). Now \( \mathbf{H} \) is a histogram array of the array \( \mathbf{Z} \) using \( m \) bins, and the \( i \)-th bin is centered at \( B_i \). Type \texttt{help hist} to become more familiar with generating histograms in Matlab.

(c) Define \( \hat{f}_Z \triangleq \mathbf{H}/(bn) \) as an empirical estimate of \( f_Z \).

5. Plot \( \hat{f}_Z \) and \( f_Z \) as functions of the array \( \mathbf{B} \). Comment on your results.
6. Now extend the definition of $\hat{f}_Z$ to the entire real line by assuming that it maintains a constant value over each bin. Note that $\hat{f}_Z$, as a function of the continuous real variable $x$, is a piece-wise constant function. Give a rough sketch of $\hat{f}_Z$ as a function of $x$.

7. Argue that

$$\int_{-\infty}^{\infty} \hat{f}_Z(x) \, dx = \sum_{i=1}^{m} \hat{f}_Z(B_i) \delta = 1,$$

and hence, $\hat{f}_Z$ is a valid probability density density

8. Give a concise summary of what you learned from this problem.
Matlab Assignment

We first calculate the cdf of $Z$. Note that $F_Z(z) = P\{Z \leq z\} = P\{g(X) \leq z\} = P\{-\mu \log(1 - X) \leq z\} = P\{\log(1 - X) \geq -z/\mu\} = P\{1 - X \geq \exp(-z/\mu)\} = P\{X \leq 1 - \exp(-z/\mu)\} = F_X(1 - \exp(-z/\mu))$.

Hence, $F_Z(z) = F_X(1 - \exp(-z/\mu))$. Note that $0 \leq 1 - \exp(-z/\mu) \leq 1$ when $z \geq 0$, and $1 - \exp(-z/\mu) < 0$ when $z < 0$. We now apply the formula for $F_X$ to $1 - \exp(-z/\mu)$ and obtain $F_Z(z) = 0$ when $z < 0$ and $F_Z(z) = (1 - \exp(-z/\mu))$ when $z \geq 0$. We then differentiate with respect to $z$ to conclude that $Z$ is an exponentially distributed random variable with parameter $\mu$.

2. See Matlab Code (below)

3. See Matlab Code (below)

4. See Matlab Code (below)

5. $\hat{f}_Z$ appears to approximate $f_Z$, so we are inclined to believe that the samples $Z_1, Z_2, \ldots, Z_{1000}$ are indeed realizations of an exponentially shaped pdf with parameter $\mu = 1$. Note the effect of changing $n$ and the bin size.

6. For $z > \max\{Z_i\}$ and $z < \min\{Z_i\}$, put $\hat{f}_z = 0$. Now for $0 \leq i \leq m$, if $B_i \leq z \leq B_{i+1}$, put $\hat{f}(z) = f(B_i)$. Now we have a function $f(z)$ which is defined on entire real line and it agrees with $H/\delta n$ at $z = B_i$, $0 \leq i \leq m$.

7. 

$$
\int_{-\infty}^{\infty} \hat{f}(z) dz = \int_{-\infty}^{\infty} \sum_{i=1}^{100} \hat{f}(B_i) [u(z - B_i) - u(z - B_{i+1})] dz
$$

$$
= \sum_{i=1}^{100} \hat{f}(B_i) \delta = \sum_{i=1}^{100} \frac{H_i}{n\delta} = \frac{1}{n} \sum_{i=1}^{100} H_i = \frac{n}{n} = 1
$$
8. This problem demonstrates that we can generate exponentially distributed random numbers by means of transforming uniformly distributed random numbers which can be easily generated on many packages such as Matlab and MathCAD.

We also learned how to generate an estimate of the pdf of an unknown r.v. given that we have access to samples of it. In particular, we used the histogram to perform such an estimation. We also emphasize that any estimate of a pdf must itself be a pdf, i.e., it must be non-negative and it must have unit area.
MATLAB Code

clear
mu=1;
m=200;
n=100000;
X=rand(n,1);
Z=-mu*log(1-X);
mx=max(Z);
mn=min(Z);
ave=mean(Z);
delta=mx/m;
B=[0:delta:mx];
H=hist(Z,B);
f_hat=H/(delta*n);
f=(1/mu)*exp(-B/mu);

subplot(2,1,1);
plot(B,f_hat,'.',B,f,'k-');
xlabel('B');
ylabel('estimated f_Z and true Z');
subplot(2,1,2);
stairs(B,f_hat);
xlabel('Z');
ylabel('estimated f_Z');
gtext('100000 samples; 200 bins')
Problem 4. Let $Z$ be a uniformly-distributed r.v. in the interval $[0, 1]$. Let
\[ f(x) = \mu^{-1}e^{-x/\mu}u(x). \]

Define a new random variable $X$ by the rule:
\[ X = g^{-1}(Z), \]
where $g^{-1}$ is the inverse function of $g$ defined by
\[ g(x) = \int_{-\infty}^{x} f(t) \, dt. \]

1. Prove that $g^{-1}$ is well defined, i.e., show that $g$ is a one-to-one mapping. You can show this fact by proving that $f(x_1) \neq f(x_2)$ whenever $x_1 \neq x_2$. (Hint: Show first that $g$ is a strictly increasing function.)
2. Derive a formula for $g^{-1}$.
3. Determine the range of the r.v. $X$.
4. Determine the p.d.f. of $X$, denoted by $f_X(x)$.

Solutions:
1. Since $f > 0$, $g(x_2) - g(x_1) = \int_{x_1}^{x_2} f(x) \, dx > 0$ as long as $x_2 > x_1$. Therefore, $g$ is strictly increasing, and hence, it is a one-to-one mapping. Therefore, $g^{-1}$ exists.

2. $g(x) = 1 - e^{-x/\mu}$. Put $z = 1 - e^{-x/\mu}$ or $x = -\mu \log(1 - z)$. Therefore, $g^{-1}(z) = -\mu \log(1 - z)$.

3. Put $X = g^{-1}(Z)$ when $0 \leq Z \leq 1$, clearly, $g^{-1}$ is an increasing function. $g^{-1}(0) = 0$ and $g^{-1}(1) = \infty$. Therefore, range of $X$ is $[0, \infty)$.

4. Use problem (1) to conclude that $f_X(x) = f(x)$. 
Problems from the text (J. Gubner). Chapter 2: Problems 21, 22, 32, 37, 40, 41.

Text, 2.21
(a) Recall that for a geometric random variable (see page 74 of the text), the probability mass function is $P\{X = n\} = (1-p)p^{n-1}, n = 1, \ldots$. Now $P\{X > n\} = \sum_{i=n+1}^{\infty}(1-p)p^i = 1 - \sum_{i=0}^{n}(1-p)p^i = 1 - (1-p)^{n+1}/(1-p) = p^n$.
(b) $P(\{X > n + k\} \mid \{X > n\}) = P(\{X > n + k\} \cap \{X > n\})/P\{X > n\} = P\{X > n + k\}/P\{X > n\} = p^{n+k}/p^n = p^k$. Interpretation: If we think of $X$ as the age-of or time-to some event (e.g., arrival time of a bus), then the fact that the event has not occurred by time $n$ does not change the probability that it will not occur for another $k$ units of time; the only thing that matters is $k$ and not $n$. In other words, the memory encapsulated in what happened till time $n$ is irrelevant to the future (i.e., from time $n$ and on). That is why we often refer to the geometric random-variable model as a memoryless model. The same is said about an exponentially distributed random variable (show this).

Moreover, knowing that the event has not occurred by time $n$, it means, in this type of random variable, that we can re-start out stop-watch at time $n$ while flushing out the past entirely as we attempt to compute probabilities of the occurrence of the event in the future. We say the process of counting the time of occurrence is \textit{regenerated} at time $n$ after knowledge that the event has not occurred by time $n$.

Text, 2.22
Let $X$ be a positive random variable with the property that $P(\{X > n + k\} \mid \{X > n\}) = P\{X > k\}$. Note that by setting $k = 1$, we have $P(\{X > n + 1\} \mid \{X > n\}) = P\{X > 1\}$. However, by hypothesis $P(\{X > n + 1\} \mid \{X > n\}) = P\{X > 1\}$. Hence, $P\{X > n + 1\} = P\{X > 1\}P\{X > n\}$. If we solve this simple recursion we obtain $P\{X > n + 1\} = P\{X > 1\}^n$; hence, we obtain $P\{X > n\} = p^n$ since according to the notation stated in the problem statement $p = P\{X > 1\}$. Next, note that $P\{X = n\} = P\{X > n - 1\} - P\{X > n\} = p^{n-1} - p^n = (1-p)p^{n-1}$ as desired.

Text, 2.32
Following the simplification for the formula for expectations of discrete random variables (see class notes), $E[X] = 2(1/3) + 5(2/3) = 4$.

Text, 2.37

Text, 2.40
$P\{X \geq 2\} = 1 - P\{X < 2\} = 1 - (e^{-0.75(0.75)}0! + e^{-0.75(0.75)}1!/1!) = 1 - e^{-0.75}[1 + 0.75] = 0.1734$. On the other hand, $E[X]/2 = 0.75/2 = 0.375$.

Text, 2.41
$E[X^2]/2 = 0.75(1 + 0.75)/4 = 0.328$, which is a bit tighter than the bound from Markov’s inequality. Using the Chernoff bound from Problem 1, we obtain 0.49, which is looser than the other two
bounds in this case (since $c$ is small). If we set $c = 10$ in this problem, then the chernoff bound gives $5.9 \times 10^{-8}$ while the Markov and Chebychev bounds yield 0.075 and 0.0131, respectively. The Chernoff bound normally performs better than the other two for rare events.