Solution of Problem Set 4

Problem 1:

a) $F_U(u) = P(U \leq u) = P(F_X(X) \leq u)$

Since $F_X(X) \in [0,1]$, then if $u < 0$, $F_U(u) = 0$ and if $u \geq 1$, $F_U(u) = 1$.

Otherwise, if $u \in [0,1)$, $F_U(u) = P\left(X \leq F_X^{-1}(u)\right) = F_X\left(F_X^{-1}(u)\right) = u$.

We used the fact that $F_X(.)$ is strictly increasing and thus its inverse is well defined. The above CDF of U is the one of uniform RV on [0,1].

Problem 2:

a) The CDF of X is given by $F_X(x) = 1 - e^{-\lambda x}, x \geq 0$. Setting $U = F_X(X) = 1 - e^{-\lambda X}$.

Therefore, we can write $X = -\frac{1}{\lambda}\log(1 - U)$.

b) The CDF of X is given by $F_X(x) = \frac{2}{\pi}\arcsin(\sqrt{x}), x \in [0,1]$.

Setting $U = F_X(X) = \frac{2}{\pi}\arcsin(\sqrt{X})$. Therefore, we can write

$X = sin^2\left(\frac{U\,\pi}{2}\right) = \frac{1}{2} - \frac{1}{2}\cos(U\pi)$.

c) The CDF of X is given by $F_X(x) = 1 - e^{-\left(\frac{x}{\alpha}\right)^{\beta}}, x \geq 0$. Setting $U = F_X(X) = 1 - e^{-\left(\frac{X}{\alpha}\right)^{\beta}}$.

Therefore, we can write $X = \alpha\,(-\log(1 - U))^{\frac{1}{\beta}}$.

Problem 3:

a) N has a geometric distribution with "success" probability $p = P\left(U \leq \frac{f(Y)}{c\,g(Y)}\right)$. The PMF of N is then given by $P(N = n) = (1-p)^{1-n}p, n \geq 1$.

To compute the value of p, we start by conditioning on Y

$$p = P\left(U \leq \frac{f(Y)}{c\,g(Y)}\right) = \int_{-\infty}^{+\infty} P\left(U \leq \frac{f(Y)}{c\,g(Y)} | Y = y\right) g(y)dy = \int_{-\infty}^{+\infty} \frac{f(y)}{c\,g(y)} \times g(y)dy = \frac{1}{c}$$

The average number of iterations required until X is successfully generated is

$$E[N] = \frac{1}{p} = c$$

b) Using Bayes theorem, we can write

$$P\left(Y \leq y \,| U \leq \frac{f(Y)}{c\,g(Y)}\right) = \frac{P\left(Y \leq y\,, U \leq \frac{f(Y)}{c\,g(Y)}\right)}{P\left(U \leq \frac{f(Y)}{c\,g(Y)}\right)} = c\,P\left(U \leq \frac{f(Y)}{c\,g(Y)}, Y \leq y\right)$$

Now, we focus on computing $P\left(U \leq \frac{f(Y)}{c\,g(Y)}, Y \leq y\right)$

$$P\left(U \leq \frac{f(Y)}{c\,g(Y)}, Y \leq y\right) = \int_{-\infty}^{y} P\left(U \leq \frac{f(Y)}{c\,g(Y)} \,|\, Y = \omega \leq y\right) g(\omega)d\omega = \int_{-\infty}^{y} \frac{f(\omega)}{c\,g(\omega)} \times g(\omega)d\omega$$

Therefore, we get

$$P\left(U \leq \frac{f(Y)}{c\,g(Y)}, Y \leq y\right) = \frac{1}{c}F(y)$$

Thus, we can write

$$P\left(Y \leq y \,| U \leq \frac{f(Y)}{c\,g(Y)}\right) = F(y)$$

Problem 4:

(a)

Recalling that

$$\frac{d}{dx}\tan^{-1}(x) = \frac{1}{1+x^2},$$

we have for $X \sim \text{Cauchy}(0,1)$ that

$$F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} \frac{1}{\pi(1+y^2)}dy = \left[\frac{1}{\pi}\tan^{-1}(x)\right]_{-\infty}^{x} = \frac{\tan^{-1}(x)}{\pi} + \frac{1}{2}.$$

(b)

The sampling method to sample $X \sim \text{Cauchy}(0,1)$ by the inversion method with $U \sim U([0,1])$ and

$$X = F^{-1}(U) = \tan\left(\pi\left(U - \frac{1}{2}\right)\right)$$

(c)

First generate an $X \sim \text{Cauchy}(0,1)$ with the above method and thereafter set $Y = \gamma X + x_0$. Then

$$F_Y(y) = \mathbb{P}(Y \le y) = \mathbb{P}\left(X \le \frac{y - x_0}{\gamma}\right) = \frac{\tan^{-1}\left(\frac{y-x_0}{\gamma}\right)}{\pi} + \frac{1}{2},$$

and

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{1}{\pi\gamma\left(1 + \left(\frac{y-x_0}{\gamma}\right)^2\right)} = f(y; x_0, \gamma).$$

Problem 5:

(a)

$$g(x) = \frac{1}{x+1},$$
$$g''(x) = \frac{2}{(1+x)^3} > 0, \quad \text{for} \quad x > 0.$$

Thus g is convex on $(0, \infty)$. We conclude

$$E\left[\frac{1}{X+1}\right] \ge \frac{1}{1+EX} \quad \text{(Jensen's inequality)}$$
$$= \frac{1}{1+10}$$
$$= \frac{1}{11}.$$

(b) If we let $h(x) = e^x, g(x) = \frac{1}{1+x}$ then h is convex and non-decreasing, and g is convex; thus by Problem 8, $e^{\frac{1}{x+1}}$ is a convex function. Therefore,

$$E\left[e^{\frac{1}{1+X}}\right] \ge e^{\frac{1}{1+EX}} \quad \text{(by Jensen's inequality)}$$
$$= e^{\frac{1}{11}}.$$

(c) If $g(x) = \ln\sqrt{x} = \frac{1}{2}\ln x$, then $g'(x) = \frac{1}{2x}$ for $x > 0$, and $g''(x) = -\frac{1}{2x^2}$. Thus g is concave on $(0, \infty)$. We conclude

$$E\left[\ln\sqrt{X}\right] = \frac{1}{2}\ln X$$
$$\le \frac{1}{2}\ln EX \quad \text{(by Jensen's inequality)}$$
$$= \frac{1}{2}\ln 10.$$

Problem 6:

1 Markov

Prob$[X \geq \alpha n]$ $p < \alpha <$

Cond. $p = \frac{1}{2}$ $\alpha = \frac{3}{4}$

$$P(X \geq \alpha n) \leq \frac{E(X)}{\alpha n} = \frac{np}{\alpha n} = \frac{p}{\alpha} = \frac{1 \cdot 4}{2 \cdot 3} = \frac{2}{3}$$

$\alpha n = \frac{3n}{4}$

2 Chebyshev

$$P(X \geq \alpha n) = P[X - np \geq \alpha n - np]$$

$$\leq P(|X - \underset{E(X)}{np}| \geq n\alpha - np)$$

$$\leq \frac{Var(X)}{(n\alpha - np)^2}$$

$$= \frac{np(1-p)}{n^2(\alpha - p)^2} = \frac{p(1-p)}{n(\alpha - p)^2}$$

$$= \frac{\frac{1}{2} \cdot \frac{1}{2}}{n\left(\frac{1}{4}\right)^2} = \frac{16}{4n} = \frac{4}{n}$$

③ chernoft:

$$P(X \geq \alpha n) \leq \min_{t \geq 0} \frac{E[e^{tX}]}{e^{\alpha n t}}$$

$$E[e^{tX}] = (pe^t + q)^n$$

$$P(X \geq \alpha n) \leq \min_{t \geq 0} \left(\frac{pe^t + q}{e^{\alpha t}}\right)^n \triangleq f(t)^n$$

To minimize $f(t)^n$, we only need to minimize $f(t)$

$$f(t) = e^{-\alpha t}(pe^t + q)$$

$$f'(t) = e^{-\alpha t}(pe^t + q)\cdot(-\alpha) + e^{-\alpha t}(pe^t) \triangleq 0$$

$$\Leftrightarrow e^t = \frac{\alpha q}{p(1-\alpha)} = 3 \quad \text{with } \alpha = \frac{3}{4},\ p = q = \frac{1}{2}$$

By checking 2nd derivative, e^t is the minimizer of $f(t)$

Then plug $e^t = \frac{\alpha q}{p(1-\alpha)} = 3$ into the inequality we can find,

$$P(X \geq \alpha n) \leq \left(\frac{16}{27}\right)^{\frac{n}{4}}$$

Problem 7:

a) We have

$$\int_0^{+\infty} f(x)\,dx = 1 \rightarrow C\int_0^{+\infty} x e^{\frac{-2}{3}x^2}\,dx = 1 \rightarrow C\left[\frac{3}{4}e^{\frac{-2}{3}x^2}\right]_0^{+\infty} = 1 \rightarrow C = \frac{4}{3}$$

b) We will a Gaussian PDF as the distribution to sample from in the acceptance rejection algorithm. More precisely, we pick $g(x) = \frac{e^{\frac{-x^2}{3}}}{\sqrt{3\pi}}$.

Now, we want to find an upper bound for the ratio, that is

$$\frac{f(x)}{g(x)} = 4\sqrt{\frac{\pi}{3}}\, x e^{\frac{-x^2}{3}} 1_{(x \geq 0)}$$

An upper bound is achieved when $x = \sqrt{\frac{3}{2}}$ and its value is $c = 2\sqrt{2\pi}\, e^{\frac{-1}{2}}$.

Algorithm:

1. Sample Y according to Gaussian with mean 0 and variance 3/2
2. Sample U a uniform RV U[0,1]
3. If $U \leq \frac{f(Y)}{c\,g(Y)} = \sqrt{\frac{2}{3}}\, Y e^{\frac{1}{2} - \frac{Y^2}{3}} 1_{(Y \geq 0)}$: accept $X = Y$, otherwise reject and return to step 1.

<u>MATLAB function:</u>

```
function X=AR(n)
i=1;
X=zeros(n,1);
while i <=n
    Y=sqrt(3/2)*randn();
    U=rand();
    if U <= sqrt(2/3)*Y*exp(1/2-Y^2/3)*(Y>=0)
        X(i) = Y;
        i = i+1;
    end
end
```

c) The CDF of X is given by $F(x)=1-e^{\frac{-2x^2}{3}}$. Thus, we can deduce that $F^{-1}(x)=\sqrt{\frac{-3}{2}\log(1-u)}$.

<u>MATLAB function:</u>

```
function X=IT(n)
X = sqrt(-3/2*log(1- rand(n,1)));
```

d) In order to obtain samples of approximately equal sizes, densities are simulated using the function AR(n*c), in order to obtain on average n samples, and exactly n realizations using the IT function (n). We obtain the plot in Figure 1. The functions AR(n*c) and IT(n) both generate realizations of the same density f. One can thus retain as relevant points of comparison the complexity of their codes, or their respective computational times. In both cases, IT(n) appears to be the best since it is both the simplest (it only requires one line of code) and the least expensive of the two functions since it does not need to reject a fraction of the simulated variables.

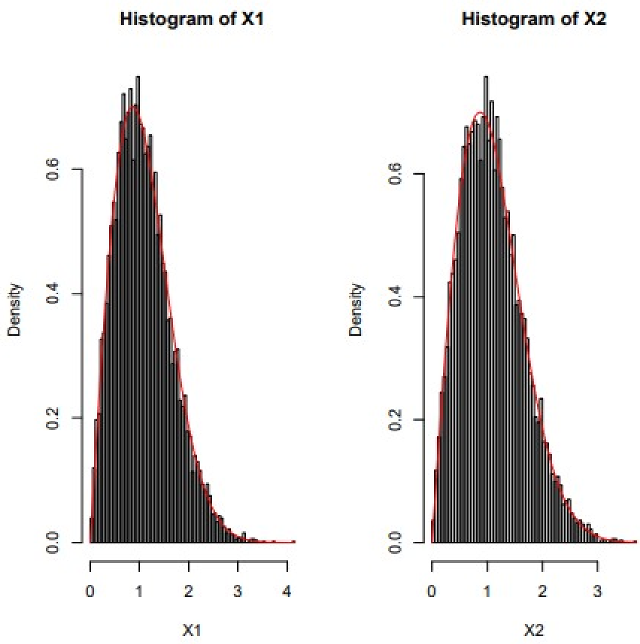


FIGURE 1 –

Problem 8:

$E[X]=1$ $\quad Var(X)=1$ $\quad E[e^{Xt}] = \frac{\lambda}{\lambda - t} = \frac{1}{1-t}$ for $t<1$

$P(X \geq a) = e^{-a}$ for $a \geq 0$

Markov : $P(X \geq a) \leq \frac{E[X]}{a} = \frac{1}{a}$

Chebyshev: $P(X \geq a) \leq P(|X - E[X]| \geq a - E[X]) \leq \frac{Var(X)}{(a-E[X])^2} = \frac{1}{(a-1)^2}$

Chernoff : $P(X \geq a) \leq \frac{E[e^{Xt}]}{e^{at}} = \frac{e^{-at}}{1-t} \triangleq f(t)$ want to minimize $f(t)$

$0<t<1$

$$f'(t) = \frac{-a(1-t)e^{-at} + e^{-at}}{(1-t)^2} \triangleq 0 \Rightarrow t = \frac{a-1}{a} \quad (a>1)$$

$$f''(t) = \frac{e^{-at}(-2-2a(-1+t) - a^2(-1+t)^2)}{(t-1)^3}$$

1° $a<1$

$$f'(t) = \frac{e^{-at}(1-a(1-t))}{(1-t)^2} > 0 \quad \text{since } t \in (0,1)$$

This means that $f(t)$ is monotonically increasing and it is minimized as $t \to 0$.

$$\lim_{t \to 0} f(t) = 1$$

$P(X \geq a) \leq 1$ (useless!)

2° $a>1$

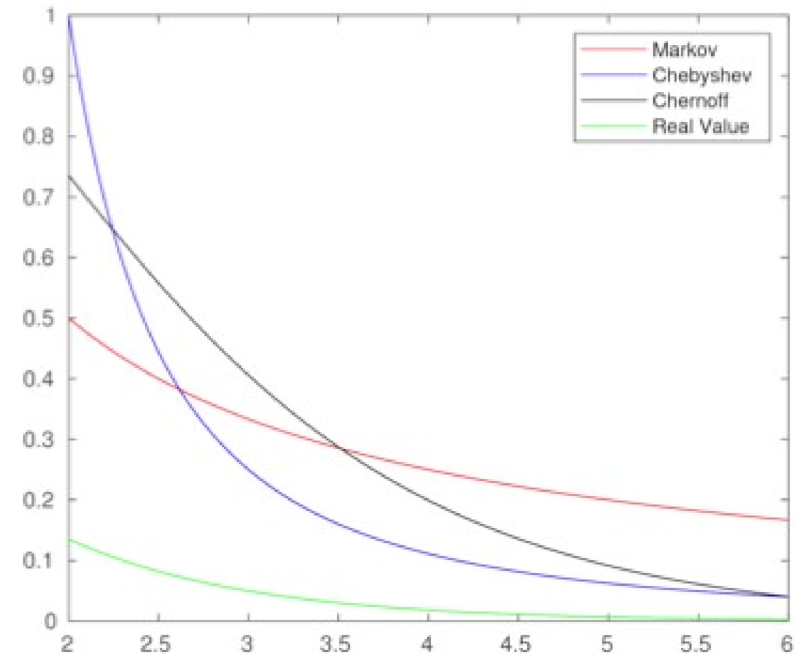
$f''(t) = a^3 e^{-(a-1)} > 0$ when $t = \frac{a-1}{a}$

Then $f(t) = a e^{1-a}$.

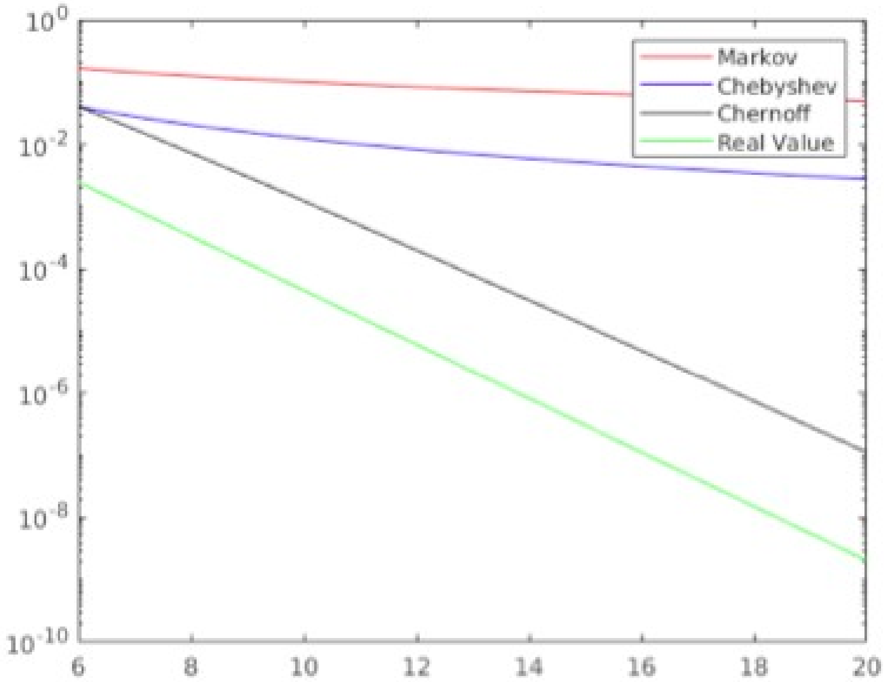
$P(X \geq a) \leq a e^{1-a}$ for $a>1$

$P(X \geq a) \leq 1$ for $0<a<1$

(a) $\frac{1}{a} < \frac{1}{(a-1)^2} \Longleftrightarrow a \in \left(\frac{3-\sqrt{5}}{2}, \frac{3+\sqrt{5}}{2}\right)$



Markov
Chebyshev
Chernoff
Real Value
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2
2.5
3
3.5
4
4.5
5
5.5
6




```
Markov=@ (a) 1./a;
Chebyshev=@(a) 1./(a-1).^2;
Chernoff=@(a) (a<1)+(a>=1).*a.*exp(1-a);
a=2:0.01:6;


figure(1)
plot(a,Markov(a),'red')
hold on
plot(a,Chebyshev(a),'blue')
plot(a,Chernoff(a),'black')
plot(a,1-expcdf(a,1),'green')
legend('Markov','Chebyshev','Chernoff','Real Value')

figure(2)
a=6:0.01:20;
semilogy(a,Markov(a),'red')
hold on
semilogy(a,Chebyshev(a),'blue')
semilogy(a,Chernoff(a),'black')
semilogy(a,1-expcdf(a,1),'green')
legend('Markov','Chebyshev','Chernoff','Real Value')
```

Problem 9:

Let A_i be the event that your friend receives an offer from the ith company, i=1,2,3,4. Then, by the union bound:

$$\begin{aligned} P\left(\bigcup_{i=1}^{4} A_i\right) &\leq \sum P(A_i) \\ &= 0.2+0.2+0.2+0.2 \\ &= 0.8 \end{aligned}$$

Thus the probability of receiving at least one offer is less than or equal to 80%.

Problem 10:

Using $\rho(X,Y) \leq 1$ and $\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$, we conclude

$$\frac{EXY - EXEY}{\sigma_X \sigma_Y} \leq 1.$$

Thus

$$\begin{aligned} EXY &\leq \sigma_X \sigma_Y + EXEY \\ &= 2 \times 1 + 2 \times 1 \\ &= 4. \end{aligned}$$

In fact, we can achieve $EXY = 4$, if we choose $Y = aX + b$.

$$Y = aX + b \quad \Rightarrow \quad \begin{cases} 2 = a + b \\ 1 = (a^2)(4) \end{cases}$$

Solving for a and b, we obtain

$$a = \frac{1}{2}, \quad b = \frac{3}{2}.$$

Problem 11:

Let F_i be the event that the ith component fails. Then,

$$\begin{aligned} P(F) &= P\left(\bigcup_{i=1}^{4} F_i\right) \\ &\leq \sum_{i=1}^{4} P(F_i) \\ &\leq \frac{4}{100} \end{aligned}$$

Problem 12:

$$
\begin{aligned}
X &\sim Geometric(p) \\
EX &= \frac{1}{p}, \\
&P(X \geq a) \\
&\leq \frac{EX}{a} \quad \text{(Using Markov's inequality)} \\
&= \frac{1}{pa}
\end{aligned}
$$

$$
\begin{aligned}
P(X \geq a) &= \sum_{k=a}^{\infty} P(X = k) \\
&= \sum_{k=a}^{\infty} q^{k-1} p \\
&= pq^{a-1} \frac{1}{1-q} \\
&= q^{a-1} \\
&= (1-p)^{a-1}
\end{aligned}
$$

We show $(1-p)^{a-1} \leq \frac{1}{pa}$ for all $a \geq 1$, $0 < p < 1$. To show this, look at the function:

$$
\begin{aligned}
f(p) &= p(1-p)^{a-1} \\
f'(p) &= 0 \quad \text{which results in} \quad p = \frac{1}{a} \\
f(p) &\leq \frac{1}{a}(1 - \frac{1}{a})^{a-1} \leq \frac{1}{a} \\
p(1-p)^{a-1} &\leq \frac{1}{a} \\
(1-p)^{a-1} &\leq \frac{1}{pa}
\end{aligned}
$$

Problem 13:

$$
\begin{aligned}
P(|X - EX| \geq 20) &\leq \frac{\mathrm{Var}(X)}{400} \\
&= \frac{225}{400} \\
&= 0.5625
\end{aligned}
$$

Problem 14:

$$\frac{X^2 + X}{2} = U$$
$$(X + \frac{1}{2})^2 - \frac{1}{4} = 2U$$
$$X + \frac{1}{2} = \sqrt{2U + \frac{1}{4}}$$
$$X = \sqrt{2U + \frac{1}{4}} - \frac{1}{2} \quad (X, U \in [0, 1])$$

By generating a random number, U, we have the desired distribution.

$$U = rand;$$
$$X = sqrt\left(2U + \frac{1}{4}\right) - \frac{1}{2};$$

Problem 15:

Using Inverse Transformation Method:

$$U - \frac{1}{2} = \frac{1}{\pi}\arctan(X)$$
$$\pi\left(U - \frac{1}{2}\right) = \arctan(X)$$
$$X = \tan\left(\pi(U - \frac{1}{2})\right)$$

Next, here is the MATLAB code:

```
U = zeros(1000, 1);
n = 100;
average = zeros(n, 1);
for   i = 1 : n
U = rand(1000, 1);
X = tan(pi * (U - 0.5));
average(i) = mean(X);
end
plot(average)
```

Cauchy distribution has no mean (Figure 12.6), or higher moments defined.

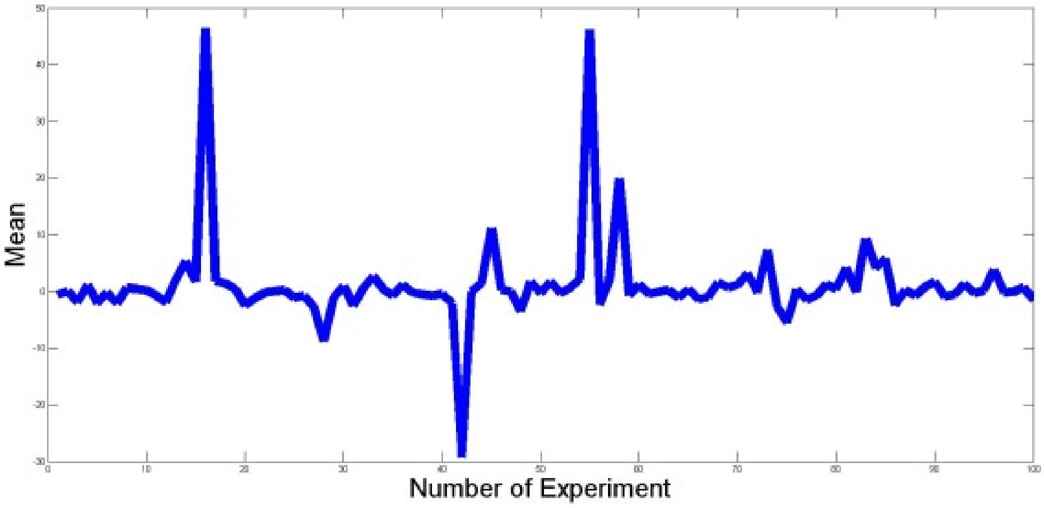


Figure : Cauchy Simulation

Problem 16:

$$f(x) = 20x(1-x)^3 \quad 0 < x < 1$$
$$g(x) = 1 \quad 0 < x < 1$$
$$\frac{f(x)}{g(x)} = 20x(1-x)^3$$

We need to find the smallest constant c such that $f(x)/g(x) \leq c$. Differentiation of this quantity yields

$$\frac{d\left(\frac{f(x)}{g(x)}\right)}{dx} = 0$$
$$\text{Thus,} \quad x = \frac{1}{4}$$
$$\text{Therefore,} \quad \frac{f(x)}{g(x)} \leq \frac{135}{64}$$
$$\text{Hence,} \quad \frac{f(x)}{cg(x)} = \frac{256}{27}x(1-x)^3$$

```
n = 1;
while(n == 1)
U1 = rand;
U2 = rand;
if   U2 <= 256/27 * U1 * (1 − U1)^3
X = U1;
n = 0;
end
end
```

Problem 17:

$$f(x) = \frac{2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad 0 < x < \infty$$

$$g(x) = e^{-x} \quad 0 < x < \infty \quad \text{(Exponential density function with mean 1)}$$

Thus, $\dfrac{f(x)}{g(x)} = \sqrt{\dfrac{2}{\pi}} e^{x - \frac{x^2}{2}}$

Thus, $x = 1$ maximizes $\dfrac{f(x)}{g(x)}$

Thus, $c = \sqrt{\dfrac{2e}{\pi}}$

$$\frac{f(x)}{cg(x)} = e^{-\frac{(x-1)^2}{2}}$$

The algorithm for generating Z is then

1. Generate Y with an exponential distribution at rate 1; that is, generate U and set $Y = -\ln(U)$.
2. Generate U
3. If $U \le e^{-(Y-1)^2/2}$, set $|Z| = Y$; otherwise go back to 1.
4. Generate U. Set $Z = |Z|$ if $U \le 0.5$, set $Z = -|Z|$ if $U > 0.5$.

Problem 18:

$Y = X/n$

$$E[Y] = E[X]/n = np/n = p$$

$$VAR[Y] = VAR[X]/n^2 = npq/n^2 = pq/n, \; q = 1 - p$$

$$P[\{|Y - p|\} > a] \le \frac{\sigma^2}{a^2} = \frac{pq}{na^2}$$

as $n \to \infty$ $P[\{|Y - p|\} > a] \to 0$ for any fixed $a > 0$

Problem 19:

$p_X(x) = 2 - 2x \quad 0 \leq x \leq 1$

$F_X(t) = \int_{-\infty}^{t} p_X(x)dx = \int_0^t (2-2x)dx = 2t - \cancel{2}.\frac{x^2}{\cancel{2}}\Big|_0^t = 2t - t^2 \quad , 0 \leq t \leq 1$

$$F_X(t) = \begin{cases} 0 & t \leq 0 \\ 2t - t^2 & 0 \leq t \leq 1 \\ 1 & t \geq 1 \end{cases} \qquad F_U(t) = \begin{cases} 0 & t \leq 0 \\ t & 0 \leq t \leq 1 \\ 1 & t \geq 1 \end{cases}$$

$F_X(t) = P(X \leq t) = P(g(U) \leq t) = P(U \leq g^{-1}(t)) = g^{-1}(t) \quad (\text{as } 0 \leq g^{-1}(t) \leq 1)$

So, $F_X(t) = g^{-1}(t) = U$

$2t - t^2 = U$

$\Leftrightarrow \quad t^2 - 2t + U = 0$

$\Delta' = 1 - U$

$t_1 = 1 - \sqrt{1-u}$ or $t_2 = 1 + \sqrt{1-u}$

$u \in [0,1] \Rightarrow t_1 \in [0,1]$ and $t_2 \in [1,2]$, so we take t_1

Hence, $t = 1 - \sqrt{1-u}$

$X = 1 - \sqrt{1-U}$

Problem 20:

We know that $X = F_Y(Y)$ is uniformily distributed on $(0,1)$. So, we can write $Y = F_Y^{-1}(X) = g(X)$, so the tranformation $g(\cdot)$ is nothing but the inverse of the CDF of Y.

⊛ We start by computing the CDF of Y. Let $t \geq 0$,

$$\begin{aligned} F_Y(t) &= \int_0^t f_Y(y)\,dy \\ &= 4a^4 \int_0^t \frac{y}{(a^2+y^2)^3}\,dy \\ &= 4a^4 \left[-\frac{1}{4(a^2+y^2)^2} \right]_0^t \\ &= 4a^4 \left[\frac{1}{4a^4} - \frac{1}{4(a^2+t^2)^2} \right] \\ &= 1 - \frac{a^4}{(a^2+t^2)^2} \end{aligned}$$

(+) To find the inverse CDF, let $x = F_Y(t) = 1 - \frac{a^4}{(a^2+t^2)^2}$

thus, $1 - x = \frac{a^4}{(a^2+t^2)^2}$

$\Rightarrow (a^2+t^2)^2 = \frac{a^4}{1-x} \quad , \quad 0 \leqslant x < 1$

$\Rightarrow a^2+t^2 = \frac{a^2}{\sqrt{1-x}} \quad , \quad 0 \leqslant x < 1$

$\Rightarrow t^2 = a^2\left[\frac{1}{\sqrt{1-x}} - 1\right] \quad , \quad 0 \leqslant x < 1$

thus, $t = a\left[\frac{1}{\sqrt{1-x}} - 1\right]^{\frac{1}{2}} \quad , \quad 0 \leqslant x < 1$

Therefore, we define $g(\cdot)$ as:

$$g(x) = a\left[\frac{1}{\sqrt{1-x}} - 1\right]^{\frac{1}{2}} \quad , \quad 0 \leqslant x < 1 .$$

Problem 21:

Let X_i be the number on the face of the die for roll i. Let X be the sum of the dice rolls. Therefore $X = \sum_{i=1}^{100} X_i$. By linearity of expectation, we write $\mathbb{E}[X] = \sum_{i=1}^{100} \mathbb{E}[X_i]$. We can compute

$$\mathbb{E}[X_i] = \sum_{j=1}^{6} j\mathbb{P}[X_i = j] = \sum_{j=1}^{6} j(1/6) = (1/6)\frac{6(7)}{2} = 7/2,$$

where we use the fact that $\sum_{j=1}^{n} j = \frac{n(n+1)}{2}$. Then we have

$$\mathbb{E}[X] = 100(7/2) = 350.$$

To use Chebyshev's inequality, the only remaining value we need to compute is the variance of X. By the independence of the dice rolls we have

$$\mathrm{Var}(X) = \mathrm{Var}\left(\sum_i X_i\right) = \sum_{i=1}^{100} \mathrm{Var}(X_i)$$

To compute the variance of a single dice roll, we use $\mathrm{Var}(X_i) = \mathbb{E}[X_i^2] + \mathbb{E}[X]^2$

$$\begin{aligned}
\mathbb{E}[X_i^2] &= \sum_{j=1}^{6} j^2\mathbb{P}[X_i = j] \\
&= \sum_{j=1}^{6} j^2(1/6) \\
&= \frac{1}{6} \cdot \frac{6(7)(13)}{6} \\
&= 91/6
\end{aligned}$$

where we use the fact that $\sum_{j=1}^{n} j^2 = \frac{n(n+1)(2n+1)}{6}$. Now we can finish computing the variance of X_i as

$$\mathrm{Var}(X_i) = \mathbb{E}[X_i^2] - \mathbb{E}[X]^2 = 91/6 - (7/2)^2 = 35/12.$$

And the variance of X is $\mathrm{Var}(X) = 100(231/4)$. Finally, we can by Chebyshev's inequality we have

$$\mathbb{P}[|X - 350| \geq 50] \leq \frac{100(35/12)}{50^2} = 7/60.$$

Problem 22:

Solution: Notice that we do not have the variance of X, so Chebyshev's bound is not applicable here. There is no upper bound on X, so Hoeffding's inequality cannot be used. We know nothing else about it's distribution so we cannot evaluate $E[e^{sX}]$ and so Chernoff bounds are not available. Since X is also not a sum of other random variables, other bounds or approximations are not available. This leaves us with just Markov's Inequality. But Markov Bound only applies on a nonnegative random variable, whereas X can take on negative values.

This suggests that we want to "shift" X somehow, so that we can apply Markov's Inequality on it. Define a random variable $Y = X + 100$, which means Y is strictly larger than 0, since X is always strictly larger than -100. Then, $E[Y] = E[X+100] = E[X] + 100 = -60 + 100 = 40$. Finally, the upper bound on X that we want can be calculated via Y, and we can now apply Markov's Inequality on Y since Y is strictly positive.

$$P(X \geq -20) = P(Y \geq 80) \leq \frac{E[Y]}{80} = \frac{40}{80} = \frac{1}{2}$$

Hence, the best upper bound on $P(X \geq -20)$ is $\frac{1}{2}$.

Problem 23:

The Chernoff bound for an arbitrary random variable X and constant c is

$$P[X \geq c] \leq \min_{s \geq 0} e^{-sc}\phi_X(s).$$

For a Poisson(α) random variable K,

$$\phi_K(s) = e^{\alpha(e^s-1)},$$

so the Chernoff bound yields

$$P[X \geq c] \leq \min_{s\geq 0} e^{-sc}e^{\alpha(e^s-1)} = \min_{s\geq 0} e^{\alpha(e^s-1)-sc}.$$

Let

$$h(s) = e^{\alpha(e^s-1)-sc}.$$

Rather than taking the derivative of $h(s)$, we can use the fact that the natural log is increasing and instead find the mimimum of $\ln h(s)$. We find that

$$\ln h(s) = \alpha e^s - \alpha - sc$$

and

$$\frac{d}{ds}\ln h(s) = \alpha e^s - c.$$

Equating this to zero and solving for s we obtain $s = \ln c/\alpha$. If $c > \alpha$ then $\ln c/\alpha > 0$ and

$$\begin{aligned} P[X \geq c] &\leq \min_{s\geq 0} e^{-cs}e^{\alpha(e^s-1)} \\ &= e^{-c\ln c/\alpha + \alpha((c/\alpha)-1)} \\ &= e^{c\ln \alpha/c + c - \alpha} \\ &= (\alpha/c)^c\, e^{c-\alpha}. \end{aligned}$$

If $c \leq \alpha$ then $\ln c/\alpha < 0$ and so we evaluate the bound at $s = 0$, in which case we obtain

$$\begin{aligned} P[X \geq c] &\leq \min_{s\geq 0} e^{-cs}e^{\alpha(e^s-1)} \\ &= 1. \end{aligned}$$

Finally

$$P[X \geq c] \leq \begin{cases} (\alpha/c)^c\, e^{c-\alpha} & c > \alpha \\ 1 & c \leq \alpha \end{cases}$$

Problem 24:

Thus the probability of receiving at least one offer is less than or equal to 80%.

2. An isolated edge in a network is an edge that connects two nodes in the network such that neither of the two nodes is connected to any other nodes in the network. Let C_n be the event that a graph randomly generated according to $G(n,p)$ model has at least one isolated edge.

   (a) Show that

$$P(C_n) \leq \binom{n}{2} p(1-p)^{2(n-2)}.$$

   (b) Show that, for any constant $b > \frac{1}{2}$, if $p = p_n = b\frac{\ln(n)}{n}$ then

$$\lim_{n\to\infty} P(C_n) = 0.$$

Solution:

There are $\binom{n}{2}$ possible edges in the graph. Let E_i be the event that the ith edge is an isolated edge, then

$$P(E_i) = p(1-p)^{2(n-2)},$$

where p in the above equation is the probability that the ith edge is present and $(1-p)^{2(n-2)}$ is the probability that no other nodes are connected to this edge. By the union bound, we have

$$\begin{aligned} P(C_n) &= P\left(\bigcup E_i\right) \\ &\leq \sum_i P(E_i) \\ &= \binom{n}{2} p(1-p)^{2(n-2)}, \end{aligned}$$

which is the desired result. Now, let $p = b\frac{\ln n}{n}$, where $b > \frac{1}{2}$.

Here, it is convenient to use the following inequality:

$$1 - x \leq e^{-x}, \quad \text{for all} \quad x \in \mathbb{R}.$$

$$P(C_n) = \binom{n}{2} p(1-p)^{2(n-2)}$$

$$= \frac{n(n-1)}{2} \frac{b\ln(n)}{n} (1-p)^{2(n-2)}$$

$$\leq \frac{n-1}{2} b\ln(n)\, e^{-2p(n-2)}$$

$$= \frac{n-1}{2} b\ln(n)\, e^{-2\cdot\frac{b\ln(n)}{n}(n-2)}$$

$$= \frac{n-1}{2} b\ln(n)\cdot n^{\frac{-2b(n-2)}{n}}$$

$$\lim_{n\to\infty} P(C_n) \leq \lim_{n\to\infty} \frac{b}{2}(n-1)\ln(n)\cdot n^{-2b}$$

$$= \lim_{n\to\infty} \frac{b}{2} \frac{(1-\frac{1}{n})\ln(n)}{n^{2b-1}}$$

$$= \lim_{n\to\infty} \frac{b}{2} \frac{\ln(n)}{n^{2b-1}}$$

Note, $f(x) = \frac{\ln x}{x^a} \to 0$ as $x \to \infty$ for all $a>0$

$$\lim_{n\to\infty} f(x) = \lim_{n\to\infty} \frac{\frac{1}{x}}{a x^{a-1}} = \lim_{n\to\infty} \frac{1}{a x^a} = 0$$

Then $\lim_{n\to\infty} P(C_n) \leq 0 \Rightarrow \lim_{n\to\infty} P(C_n) = 0$.

Problem 25:

$$f(x) = \frac{4}{3\sqrt{\pi}} x^{\frac{3}{2}} e^{-x}, x > 0$$
$$g(x) = \frac{2}{5} e^{-\frac{2x}{5}} \quad x > 0$$
$$\frac{f(x)}{g(x)} = \frac{10}{3\sqrt{\pi}} x^{\frac{3}{2}} e^{-\frac{3x}{5}}$$
$$\frac{d\left(\frac{f(x)}{g(x)}\right)}{dx} = 0$$
$$\text{Hence,} \quad x = \frac{5}{2}$$
$$c = \frac{10}{3\sqrt{\pi}} \left(\frac{5}{2}\right)^{\frac{3}{2}} e^{\frac{-3}{2}}$$
$$\frac{f(x)}{cg(x)} = \frac{x^{\frac{3}{2}} e^{\frac{-3x}{5}}}{\left(\frac{5}{2}\right)^{\frac{3}{2}} e^{\frac{-3}{2}}}$$

We know how to generate an Exponential random variable.

- Generate a random number U_1 and set $Y = -\frac{5}{2} \log U_1$.
- Generate a random number U_2.
- If $U_2 < \frac{Y^{\frac{3}{2}} e^{\frac{-3Y}{5}}}{\left(\frac{5}{2}\right)^{\frac{3}{2}} e^{\frac{-3}{2}}}$, set $X = Y$. Otherwise, execute the step 1.