1. $E[X] = \mathbb{P}(0 < Z < \beta) = \beta$

$E[Y] = \mathbb{P}(\alpha < Z < 1) = 1 - \alpha$

$E[XY] = \mathbb{P}(X=1, Y=1) = \mathbb{P}(0 < Z < \beta, \alpha < Z < 1)$

$= \mathbb{P}(\alpha < Z < \beta) = \beta - \alpha$

$\mathrm{Cov}(X,Y) = E[XY] - E[X]E[Y] = \beta - \alpha - \beta(1-\alpha)$

$= \beta - \alpha - \beta + \beta\alpha$

$= \beta\alpha - \alpha$

2. $X \sim N(\mu, \sigma^2)$

$Y \sim N(\mu, \sigma^2)$

$X + Y \sim N(2\mu, 2\sigma^2)$

$X - Y \sim N(0, 2\sigma^2)$

For a normal distribution, the MGF is given by $e^{\mu t + \frac{1}{2}\sigma^2 t^2}$

$$M_{X+Y}(t) = e^{2\mu t + \sigma^2 t^2}$$

$$M_{X-Y}(t) = e^{\sigma^2 t^2}$$

The bivariate random vector $(X+Y, X-Y)$ has a MGF given by

$$M_{(X+Y, X-Y)}(t_1, t_2) = E\left[e^{t_1(X+Y) + t_2(X-Y)}\right] = E\left[e^{t_1 X + t_1 Y + t_2 X - t_2 Y}\right]$$

$$= E\left[e^{X(t_1+t_2)} e^{Y(t_1-t_2)}\right]$$

$$= E\left[e^{X(t_1+t_2)}\right] E\left[e^{Y(t_1-t_2)}\right] \quad \text{since } X \perp\!\!\!\perp Y$$

$$= M_X(t_1+t_2) M_Y(t_1-t_2)$$

$$= e^{\mu(t_1+t_2) + \frac{1}{2}\sigma^2(t_1+t_2)^2} e^{\mu(t_1-t_2) + \frac{1}{2}\sigma^2(t_1-t_2)^2}$$

$$= e^{\mu t_1 + \mu t_2 + \mu t_1 - \mu t_2 + \frac{1}{2}\sigma^2(t_1^2 + 2t_1t_2 + t_2^2 + t_1^2 - 2t_1t_2 + t_2^2)}$$

$$= e^{2\mu t_1 + \frac{1}{2}\sigma^2(2t_1^2 + 2t_2^2)} = e^{2\mu t_1 + \sigma^2 t_1^2 + \sigma^2 t_2^2}$$

$$= e^{2\mu t_1 + \sigma^2 t_1^2} e^{\sigma^2 t_2^2} = M_{X+Y}(t_1) M_{X-Y}(t_2)$$

Since the bivariate MGF is equal to a product of its individual MGFs, $X+Y$ and $X-Y$ are independent.

3. a) $Y = ZX$

$$F_Y(x) = P(Y \le x) = P(ZX \le x)$$
$$= P(X \le x)P(Z=1) + P(X \ge -x)P(Z=-1)$$
$$= \tfrac{1}{2}P(X \le x) + \tfrac{1}{2}P(X \ge -x)$$

X is normal, so it is symmetric and $P(X \le x) = P(X \ge -x)$.

$$F_Y(x) = P(X \le x)$$
$$= F_X(x)$$

Since X and Y have equivalent CDFs, Y is normally distributed

b) $E[X] = 0$ since $X \sim N(0,1)$

$E[Y] = 0$ since $Y \sim N(0,1)$

$$E[XY] = E[X(ZX)] = E[X^2 Z]$$
$$= E[X^2]E[Z] \quad \text{since } X \perp\!\!\!\perp Z$$

$E[Z] = \frac{1}{2}(1) + \frac{1}{2}(-1) = 0$, so

$$E[XY] = E[X^2]E[Z] = 0$$

$$\mathrm{Cov}(X,Y) = E[XY] - E[X]E[Y]$$
$$= 0 - 0$$
$$= 0$$

Since $\mathrm{cov}(X,Y) = 0$, X and Y are uncorrelated.

c) $X \sim N(0,1)$

$Y \sim N(0,1)$

$$\left.\begin{aligned} P(X>1) &= 1 - F_X(1) = 1 - 0.841 = 0.159 \\ P(Y>1) &= 1 - F_Y(1) = 1 - 0.841 = 0.159 \end{aligned}\right\} \text{using normcdf calculator}$$

$$P(X>1, Y>1) = P(X>1, XZ>1) = P(X>1, Z=1)$$
$$= P(X>1)P(Z=1) \quad \text{since } X \perp\!\!\!\perp Z$$
$$= 0.159\left(\tfrac{1}{2}\right)$$

$$P(X>1, Y>1) = 0.159\left(\tfrac{1}{2}\right) \neq 0.159(0.159) = P(X>1)P(Y>1)$$

Since $P(X>1, Y>1) \neq P(X>1)P(Y>1)$, X and Y are not independent.

4. a) $E[S_n] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \lambda = n\lambda$

$$P(S_n \geq m) \leq \frac{E[S_n]}{m} = \frac{n\lambda}{m}$$

b) $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{S_n}{n}$

$$P(S_n \geq m) = P(\bar{X}_n n \geq m) = P\left(\bar{X}_n \geq \frac{m}{n}\right)$$

$$= P\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \geq \frac{\frac{m}{n} - \mu}{\sigma/\sqrt{n}}\right)$$

$$= P\left(\frac{\bar{X}_n - \lambda}{\sqrt{\lambda/n}} \geq \frac{\frac{m}{n} - \lambda}{\sqrt{\lambda/n}}\right) \quad \text{since } V[X_i] = \lambda,\ \sigma^2 = \lambda \Rightarrow \sigma = \sqrt{\lambda}$$

$$= 1 - P\left(\frac{\bar{X}_n - \lambda}{\sqrt{\lambda/n}} \leq \frac{\frac{m}{n} - \lambda}{\sqrt{\lambda/n}}\right)$$

$$= 1 - \Phi\left(\frac{\frac{m}{n} - \lambda}{\sqrt{\lambda/n}}\right) \quad \text{as } n \to \infty \text{ by CLT}$$

$$= \Phi\left(-\frac{\frac{m}{n} - \lambda}{\sqrt{\lambda/n}}\right)$$

c) See script

```
lambda = 1;
n = 100;
m = 120;
N = 10000;

% calculate Markov bound
markov_bound = n * lambda / m;

% calculate CLT estimate
clt_estimate = normcdf(-(m / n - lambda) / (sqrt(lambda / n)));

% simulate probability that S_n >= m
counter = 0;
for i=1:N
    S_n = sum(poissrnd(lambda, n, 1));
    if (S_n >= m)
        counter = counter + 1;
    end
end

probability = counter / N;
fprintf("P(S_n >= m): %.3f\n", probability);
fprintf("Markov bound: %.3f\n", markov_bound);
fprintf("CLT estimate: %.3f\n", clt_estimate);

% Through these values, it is evident that the Markov bound is not tight,
% and that the CLT estimate is a decently accurate estimate of the true
% probability.
```

```
P(S_n >= m): 0.031
Markov bound: 0.833
CLT estimate: 0.023
```

Published with MATLAB® R2023b

5. a) $I = \int_0^1 x^3 dx = \int_0^1 x^3 \frac{1}{1-0} dx = E[X^3],\ X \sim U[0,1]$

See code for implementation of samples.

b) $I = \int_0^1 x^3 dx = \int_0^1 \frac{x^3}{f(x|\alpha, \beta)} f(x|\alpha,\beta)\, dx$ $\quad \alpha = 4,\ \beta = 1$

$$= E_f\left[\frac{x^3}{f(x|4,1)}\right]$$

See code for implementation.

$$f(x|4,1) = \frac{(4+1-1)!}{(4-1)!(1-1)!} x^{4-1}(1-x)^{1-1}$$

$$= \frac{4!}{3!} x^3 = 4x^3$$

$$E_f\left[\frac{x^3}{f(x|4,1)}\right] = E_f\left[\frac{x^3}{4x^3}\right] = \frac{1}{4}$$

In our code, we got an estimate of $0.2\overline{5}$ with a single sample. This makes sense because betapdf gives $4x^3$ for $\alpha = 4$ and $\beta = 1$, which causes the x^3 to cancel out leaving us with $1/4$.

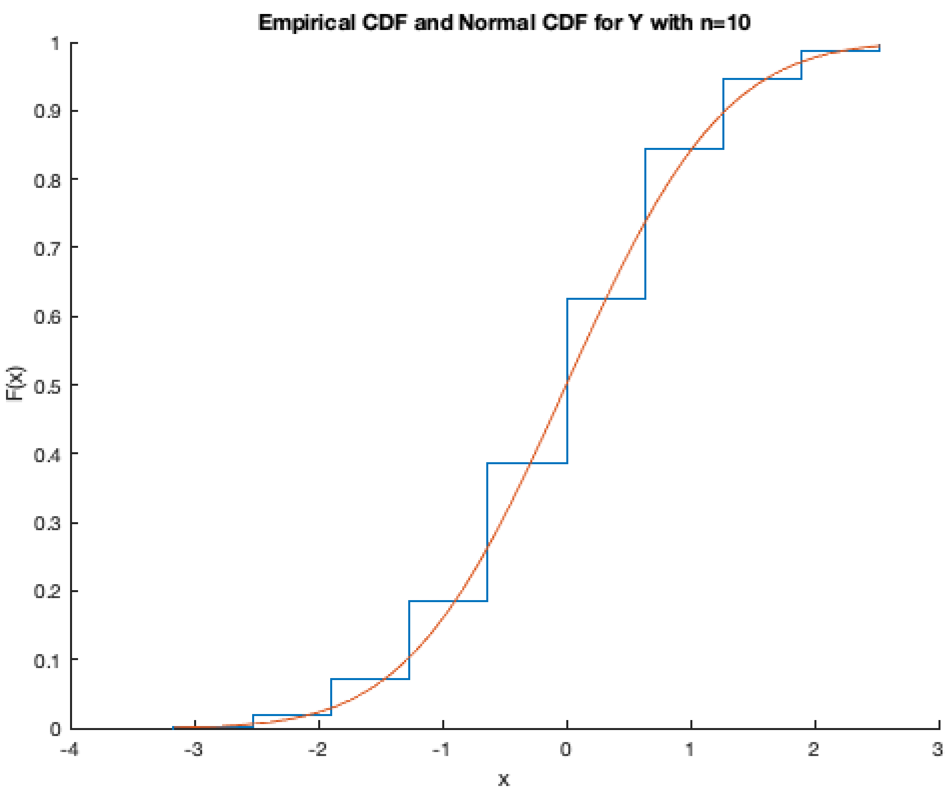
```
% part a
N = 100;
samples = rand(100, 1);
I_estimate = mean(samples .^ 3);
fprintf("Estimate of I with uniform distribution: %.3f\n", I_estimate);

% part b
alpha = 4;
beta = 1;
x = betarnd(alpha, beta); % only one sample
beta_estimate = x^3 / betapdf(x, alpha, beta);
fprintf("Estimate of I with beta distribution: %.3f\n", beta_estimate);

% See the notebook for the explanation.
```

```
Estimate of I with uniform distribution: 0.261
Estimate of I with beta distribution: 0.250
```

Published with MATLAB® R2023b

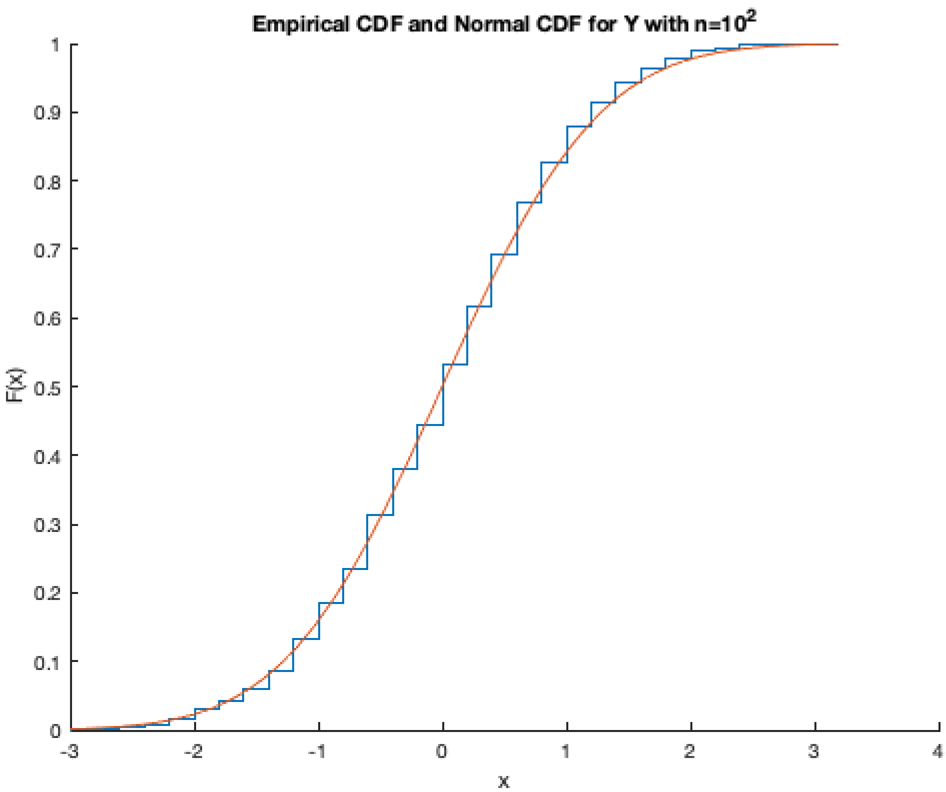
6. See my code.

```
% part a
n = 10;
N = 1000;
Y = zeros(N, 1);
for i=1:N
    X = (rand(n, 1) < 0.5) * 2 - 1;
    Y(i) = mean(X) * sqrt(n);
end

% part b
hold on
ecdf(Y);
fplot(@(x) normcdf(x), [min(Y), max(Y)]);
title("Empirical CDF and Normal CDF for Y with n=10");
xlabel("x");
ylabel("F(x)");
hold off
```



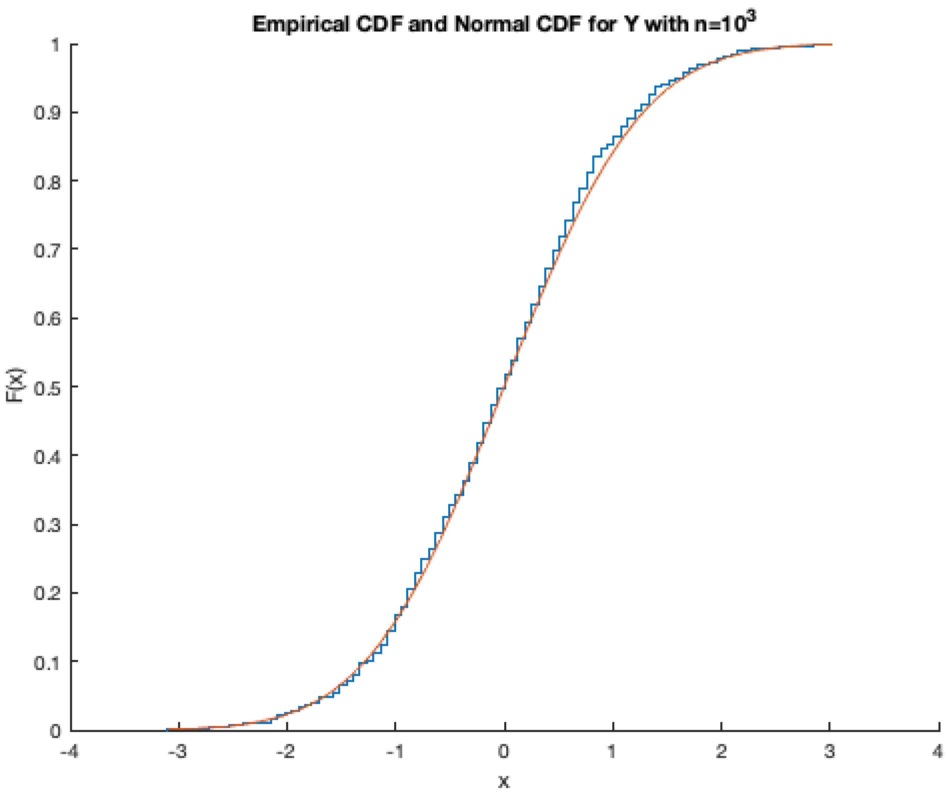
Published with MATLAB® R2023b

```
% plot for n = 10^2
n = 10^2;
N = 1000;
Y = zeros(N, 1);
for i=1:N
    X = (rand(n, 1) < 0.5) * 2 - 1;
    Y(i) = mean(X) * sqrt(n);
end

hold on
ecdf(Y);
fplot(@(x) normcdf(x), [min(Y), max(Y)]);
title("Empirical CDF and Normal CDF for Y with n=10^2");
xlabel("x");
ylabel("F(x)");
hold off
```



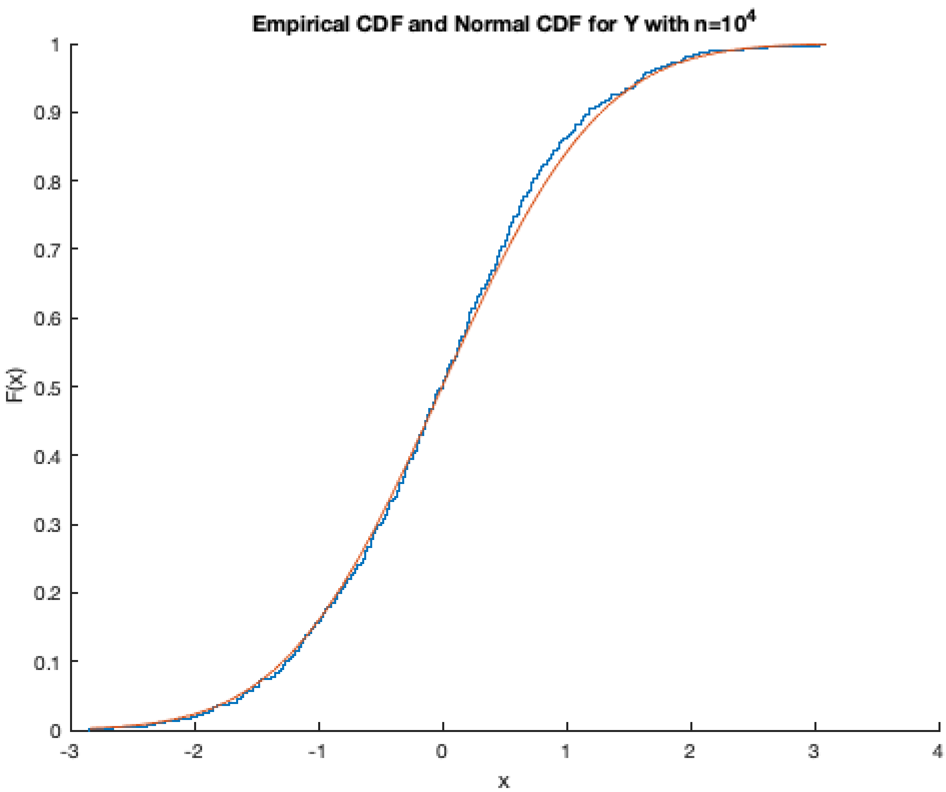
Published with MATLAB® R2023b

```
% plot for n = 10^3
n = 10^3;
N = 1000;
Y = zeros(N, 1);
for i=1:N
    X = (rand(n, 1) < 0.5) * 2 - 1;
    Y(i) = mean(X) * sqrt(n);
end

hold on
ecdf(Y);
fplot(@(x) normcdf(x), [min(Y), max(Y)]);
title("Empirical CDF and Normal CDF for Y with n=10^3");
xlabel("x");
ylabel("F(x)");
hold off
```



Published with MATLAB® R2023b

```
% plot for n = 10^4
n = 10^4;
N = 1000;
Y = zeros(N, 1);
for i=1:N
    X = (rand(n, 1) < 0.5) * 2 - 1;
    Y(i) = mean(X) * sqrt(n);
end

hold on
ecdf(Y);
fplot(@(x) normcdf(x), [min(Y), max(Y)]);
title("Empirical CDF and Normal CDF for Y with n=10^4");
xlabel("x");
ylabel("F(x)");
hold off
```



Published with MATLAB® R2023b