

HW1

Jieqi Tu

Question 1

(a)

$$\int_0^{\infty} x(1+x^2)^{-2} dx$$

Set $u = (1+x^2)$, then we can get: $\int_0^{\infty} x(1+x^2)^{-2} dx = \frac{1}{2} \int_1^{\infty} u^{-2} du = \frac{1}{2}$

In order to transform \int_0^{∞} to \int_0^1 , set $y = \frac{1}{1+x}$.

So $dx = -\frac{1}{y^2} dy$. Then we have $\int_1^0 -\frac{1}{y^2} \frac{(1/y-1)}{(1+(1/y-1)^2)^2} dy$.

```
# set seed
set.seed(1029)

# check theoretical value of the given integral using R:
integral_a = function(x) {
  x*(1+x^{2})^{-2}
}
integrate(integral_a, lower = 0, upper = Inf)$value
```

```
## [1] 0.5
```

```
# simulate using uniform distribution
s = runif(100000)
transform = (1/s^2)*(1/s-1)/(1+(1/s-1)^2)^2
mean(transform)
```

```
## [1] 0.4983382
```

(b)

$$\int_{-\infty}^{\infty} e^{-x^2} dx$$

Because we already know the probability density function of Normal distribution: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}}$, we can let $\mu = 1$, $\sigma = 1/\sqrt{2}$. Then according to the property of probability density function, we have: $\int_{-\infty}^{\infty} e^{x^2} dx = \sqrt{\pi}$

To transform upper bound and lower bound, let $y = \frac{e^x}{1+e^x}$. Then we have $\int_0^1 e^{\frac{\log it^2(y)}{y(1-y)}} dy$

```
# set seed
set.seed(1029)

# check theoretical value of the given integral using R:
integral_b = function(x) {
  exp(-x^2)
}
integrate(integral_b, lower = -Inf, upper = Inf)$value
```

```
## [1] 1.772454
```

```
# simulate using uniform distribution
s = runif(100000)
transform = exp(-(log(s/(1-s)))^2)/(s*(1-s))
mean(transform)
```

```
## [1] 1.764829
```

(c)

$$\int_2^4 \ln x dx$$

$\int_2^4 \ln x dx = x \ln x \Big|_2^4 - \int_2^4 x d \ln x = 4 \ln 4 - 2 \ln 2 - 2 \approx 2.1588$. To transform the upper and lower bound, set $y = \frac{x-2}{2}$. Then we have $\int_0^1 2 \ln(2y+2) dy$.

```
# set seed
set.seed(1029)

# check theoretical value of the given integral using R:
integral_c = function(x) {
  log(x)
}
integrate(integral_c, lower = 2, upper = 4)$value
```

```
## [1] 2.158883
```

```
# simulate using uniform distribution
s = runif(1000000000)
transform = 2*log(2*s+2)
mean(transform)
```

```
## [1] 2.158887
```

(d)

$$\int_0^1 \int_0^1 e^{x+y} dx dy$$

The exact value of this integral is $\int_0^1 e^x dx \int_0^1 e^y dy = (e-1)^2 \approx 2.952492$.

```
# set seed
set.seed(1029)

# simulate using uniform distribution
x = runif(100000)
y = runif(100000)
transform = exp(x+y)
mean(transform)
```

```
## [1] 2.950352
```

Question 2

$$\text{Cov}(U, \sqrt{1-U^2}) = E(U^2\sqrt{1-U^2}) - E(U^2)E(\sqrt{1-U^2}).$$

$$E(U^2) = \int_0^1 u^2 du = u/3|_0^1 = \frac{1}{3}.$$

$$\text{Let } u = \sin\theta, du/d\theta = \cos\theta. \text{ Then } E(U^2\sqrt{1-U^2}) = \int_0^{\pi/2} \sin^2\theta \cos^2\theta d\theta = \int_0^{\pi/2} \frac{1}{4} \sin^2 2\theta d\theta = \frac{1}{8} \int_0^{\pi/2} 1 - \cos 4\theta d\theta = \frac{\pi}{16}$$

$$E(\sqrt{1-U^2}) = \int_0^1 \sqrt{1-u^2} du = \int_0^{\pi/2} \cos\theta d\sin\theta = \int_0^{\pi/2} \cos^2\theta d\theta = \int_0^{\pi/2} \frac{1+\cos 2\theta}{2} d\theta =$$

$$\int_0^{\pi/2} (\frac{1}{2}\theta + \frac{1}{4}\sin 2\theta) d\theta = \frac{\pi}{4}$$

$$\text{Cov}(U^2, \sqrt{1-U^2}) = \frac{\pi}{16} - \frac{1}{3} \times \frac{\pi}{4} \approx -0.06544985$$

```
# set seed
set.seed(1029)

# simulate
u = runif(1000000)
y = u^2
z = sqrt(1 - y)
cov(y, z)
```

```
## [1] -0.06543874
```

Question 3

```
# set seed
set.seed(1029)

# simulate
sim_n = 10000
n_value = c()
for(i in 1:sim_n) {
  count = 0
  sum = 0
  while(sum <= 1) {
    u = runif(1)
    sum = sum + u
    count = count + 1
  }
  n_value[i] = count
}
```

```
}  
  
mean(n_value)
```

```
## [1] 2.7068
```

Question 4

```
# set seed  
set.seed(1029)  
  
# simulate  
n_sim = 100000  
n_value = c()  
for (i in 1:n_sim) {  
  count = 0  
  prod = 1  
  while(prod >=exp(-3)) {  
    u = runif(1)  
    prod = prod * u  
    count = count + 1  
  }  
  n_value[i] = count  
}  
mean(n_value)
```

```
## [1] 4.00038
```

Question 5

```
library(extraDistr)
```

```
##  
## Attaching package: 'extraDistr'  
  
## The following object is masked from 'package:purrr':  
##  
##      rdunif
```

```
set.seed(1029)  
n_sim = 100000  
n_list = numeric(n_sim) # to store all the values of the number of experiments  
for (i in 1:n_sim) {  
  d_list = c()  
  statement = F  
  n_value = 0
```

```

while (statement == F) {
  sum = 0
  sum12 = sum(rdunif(2, 1, 6))
  n_value = n_value + 1
  d_list = c(d_list, sum12)
  for (j in 2:12) {
    # whether or not this outcome shows up in our experiments
    within = as.numeric(j %in% d_list)
    sum = sum + within
  }
  # If one outcome is not present in the experiments, we will start a new experiment
  statement = (sum == 11)
}
n_list[i] = n_value
}
round(mean(n_list))

```

```
## [1] 61
```

Question 6

(a)

Since $p_j = \frac{(j-1)!}{(j-r)!(r-1)!} p^r (1-p)^{j-r}$, $p_{j+1} = \frac{(j)!}{(j+1-r)!(r-1)!} p^r (1-p)^{j+1-r}$, we have $\frac{p_{j+1}}{p_j} = j(1-p)/(j+1-r)$

(b)

The algorithm to generate a negative binomial variable:

- Generate a random variable U , which follows uniform distribution. $U \sim Uniform(0,1)$. In this case, r means the number of success in a total number of j trials. However, the last trial is defined to be successful.
- Set $c = 1 - p$, $j = r$, $pr = p^r$, and $F = pr$.
- If $U < F$, stop and report j .
- Then update: $pr = \frac{c*pr}{j+1-r}$, $F = F + pr$, $j = j + 1$.
- Go to step 3 and repeat.

(c)

Theoretical value calculation: $E(X) = \frac{rp}{1-p} = \frac{5}{0.6} = 12.5$, $Var(X) = \frac{r(1-p)}{p^2} = \frac{5*0.6}{0.16} = 18.75$

```

# set seed
set.seed(1029)

# simulate
n_sim = 1000000
p = 0.4
r = 5
c = 1 - p
n_value = c()
for (i in 1:n_sim) {
  j = r

```

```

pr = p^r
u = runif(1, 0, 1)
F_store = pr
while (u>F_store) {
  pr = c*pr*j/(j+1-r)
  F_store = F_store + pr
  j = j + 1
}
n_value[i] = j
}

mean(n_value)

```

```
## [1] 12.50136
```

```
var(n_value)
```

```
## [1] 18.73096
```

Question 7

Brute force to solve $E(|Z|)$:

$E(|Z|) = 2 \int_0^\infty z \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz =$. Let $u = e^{-\frac{1}{2}z^2}$, then we have $du/dz = -ze^{-\frac{1}{2}z^2}$. So $E(|Z|) = -\int_1^0 \frac{2}{\sqrt{2\pi}} du = \frac{2}{\sqrt{2\pi}} \approx 0.798$

```

set.seed(1029)

# create a large number of standard normally distributed variables
z = rnorm(10000000, 0, 1)
abs_z = abs(z)

# check the average
mean(abs_z)

```

```
## [1] 0.7980856
```

Question 8

$E(X) = \int_0^1 x \frac{e^x}{e-1} dx = \frac{1}{e-1} e^x (x-1) \Big|_0^1 \approx 0.581977$.

$Var(X) = E(X^2) - (E(X))^2 = \frac{1}{e-1} \int_0^1 x^2 e^x dx - (\frac{1}{e-1})^2 = \frac{e-2}{e-1} - (\frac{1}{e-1})^2 \approx 0.079326$ Simulation algorithm:

Let $F(X)$ be the CDF for x , so $F(X) = \frac{1}{e-1} \int_0^x e^t dt = \frac{1}{e-1} (e^x - 1)$.

Let $U = F(X)$, then $U \sim Uniform(0, 1)$. So we have $x = \log[(e-1)u + 1]$.

```

# set.seed
set.seed(1029)

# simulate
n_sim = 100000
x_value = c()

```

```
for(i in 1:n_sim) {  
  # create the CDF for X  
  u = runif(1, 0, 1)  
  # express X using CDF  
  x = log((exp(1)-1)*u+1)  
  x_value[i] = x  
}  
  
mean(x_value)
```

```
## [1] 0.5820019
```

```
var(x_value)
```

```
## [1] 0.07966918
```