In [2]:
```python
import random as rnd

import matplotlib.pyplot as plt

import numpy as np

import math

from scipy.stats import norm

from PIL import Image
from IPython.display import display
```

Q1

Random Digit

Posted on October 3, 2012 by Jonathan Mattingly | Comments Off

Let D_i be a random digit chosen uniformly from $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Assume that each of the D_i are independent.

Let X_i be the last digit of D_i^2. So if $D_i = 9$ then $D_i^2 = 81$ and $X_i = 1$. Define $\bar{X}_n$ by

$$\bar{X}_n = \frac{X_1 + \cdots + X_n}{n}$$

1. Predict the value of $\bar{X}_n$ when n is large.
2. Find the number ϵ such that for $n = 10,000$ the chance that you prediction is off by more than ϵ is about 1/200.
3. Find approximately the least value of n such that your prediction of $\bar{X}_n$ is correct to within 0.01 with probability at least 0.99 .
4. If you just had to predict the first digit of $\bar{X}_{100}$, what digit should you choose to maximize your chance of being correct, and what is that chance ?

[Pitman p206, #30]

$$x_i = \{0, 1, 4, 9, 6, 5, 6, 9, 4, 1\}$$

a) $\bar{X}_n = \mathbb{E}(X_i) = \sum x \, \mathbb{P}(X = x)$

x	0	1	4	5	6	9
$\mathbb{P}(X=x)$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{2}{10}$

$$\therefore \mathbb{E}(X_i) = 0 \cdot \frac{1}{10} + 1 \cdot \frac{2}{10} + 4 \cdot \frac{2}{10} + 5 \cdot \frac{1}{10} + 6 \cdot \frac{2}{10} + 9 \cdot \frac{2}{10}$$

$$= \frac{2}{10} + \frac{8}{10} + \frac{5}{10} + \frac{12}{10} + \frac{18}{10}$$

$$= \frac{20}{10} + \frac{20}{10} + \frac{5}{10} = \frac{45}{10} = \frac{9}{2}$$

b) find ε such that for $n = 10\,000$

$$\mathbb{P}\left(\left| \bar{X}_{obs} - \bar{X}_n \right| \geq \varepsilon \right) = \frac{1}{200}$$

$$\mathbb{P}(a < N_n < b) = \int_a^b \phi_{\mu,\sigma}(x)\,dx$$

$$\mathbb{P}\left(\underbrace{\frac{a-\mu}{\sigma}}_{A} < \frac{N_n - \mu}{\sigma} < \underbrace{\frac{b-\mu}{\sigma}}_{B} \right) = \int_A^B \phi_{0,1}(z)\,dz$$

want $\mathbb{P}(-\varepsilon < N_n - np < \varepsilon) = \int_{-\frac{\varepsilon}{\sigma}}^{\frac{\varepsilon}{\sigma}} \phi_{0,1}(z)\,dz$

$-\varepsilon = A\sigma \Rightarrow A = \frac{-\varepsilon}{\sigma}$

$$= \Phi\left(\frac{\varepsilon}{\sigma}\right) - \Phi\left(-\frac{\varepsilon}{\sigma}\right)$$

$$= 2\Phi\left(\frac{\varepsilon}{\underbrace{\sigma}_{SD(\bar{x})}}\right) - 1$$

Square root law $\Rightarrow 2\Phi\left(\frac{\varepsilon\sqrt{n}}{\underbrace{\sigma}_{SD(x)}}\right) - 1$

need to ask prof a bit more abt this

$$\therefore\ 2\Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) - 1 = \frac{1}{200} \Rightarrow \Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) = \frac{199}{200}$$

$\Rightarrow \underline{\underline{0 \cdot 085}}$

$$\mathbb{E}(x_i) = \frac{9}{2}$$

$$\mathrm{Var}(x_i) = \mathbb{E}(x_i^2) - (\mathbb{E}(x_i))^2 = 9 \cdot 05$$

c) $\mathbb{P}(|\bar{X}_n - 4.5| > 0.1) < 0.01 = 2\left(1 - \Phi\left(\varepsilon \frac{\sqrt{n}}{\sigma}\right)\right)$

$$\Rightarrow \Phi^{-1}(0.995) = 2.58 \Rightarrow n = \left(\frac{2.58 \cdot 3.008}{0.01}\right)^2$$

$$\Rightarrow \underline{\underline{n \approx 6 \cdot 10^5}}$$

d) $\bar{X}_{100} \approx 4.5$

$$\therefore \mathbb{P}(4 \le \bar{X}_{100} < 5) \to \mathbb{P}(|\bar{X}_n - 4.5| > 0.5) = 2\left(1 - \Phi\left(0.5 \frac{\sqrt{n}}{\sigma}\right)\right)$$

$$\Rightarrow \mathbb{P}(|\bar{X}_n - 4.5| \ge 0.5) = \underline{\underline{0.9}}$$

Q2

Change of Variable: Gaussian

Posted on October 17, 2012 by Jonathan Mattingly | Comments Off

Let Z be a standard Normal random variable (ie with distribution $N(0,1)$). Find the formula for the density of each of the following random variables.

1. $3Z+5$
2. $|Z|$
3. Z^2
4. $\frac{1}{Z}$
5. $\frac{1}{Z^2}$

[based on Pitman p. 310, #10]

1) $\mathbb{E}(3Z+5) = \mathbb{E}(3Z) + \mathbb{E}(5) = 3\mathbb{E}(Z) + 5$

$$= 0 + 5 = 5$$

$$\text{Var}(3Z+5) = 3^2 \cdot \text{Var}(Z) = 9\text{Var}(Z) = 9.$$

$$\Rightarrow \sigma = 3$$

$$\therefore \text{density} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$= \underline{\underline{\frac{1}{3\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-5}{3}\right)^2}}}$$

3) if X has density $f_X(x)$ strictly increasing / decreasing. (one-to-one)

& $Y = g(X)$ then range $Y = g(a) \longleftrightarrow g(b)$

then $f_Y(y) = \dfrac{f_X(x)}{\left|\frac{dy}{dx}\right|}$ where $y = g(x)$

$\therefore$ Z has $f_Z(z)$

$Y = Z^2$ $\therefore y = z^2 \Rightarrow \frac{dy}{dz} = 2z$

$z = \pm\sqrt{y}$ $\therefore$ many-to-one

$$f_Y(y) = \sum_{\{z:\, g(z)=y\}} \frac{f_Z(z)}{\left|\frac{dy}{dz}\right|}$$

$$= \sum_{z:\pm\sqrt{y}} \frac{f_Z(z)}{\frac{dy}{dz}} = \frac{f_Z(\sqrt{y})}{|2\sqrt{y}|} + \frac{f_Z(-\sqrt{y})}{|-2\sqrt{y}|}$$

$$= \frac{1}{2\sqrt{y}}\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y} + \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y}\right) = \frac{1}{\sqrt{2\pi y}}e^{-\frac{1}{2}y}$$

5) $Y = \frac{1}{Z^2}$

$$f_Y(y) = \sum_{y} \frac{f_Z(z)}{\left|\frac{dy}{dz}\right|}$$ $\frac{dy}{dz} = -2 \cdot z^{-3}$

$\frac{1}{\left|\frac{dy}{dz}\right|} = \frac{1}{\frac{2}{z^3}} = \frac{z^3}{2}.$

$$= \frac{z^3}{2}\left(\phi\left(\sqrt{\tfrac{1}{y}}\right) + \phi\left(-\sqrt{\tfrac{1}{y}}\right)\right)$$ $y = \frac{1}{z^2} \Rightarrow z^2 = \frac{1}{y}$

$\Rightarrow z = \pm\sqrt{\frac{1}{y}}$

$$= \frac{1}{2}\left(\sqrt{\tfrac{1}{y}}\right)^3\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2y}} + \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2y}}\right)$$

$$= \frac{1}{\sqrt{2\pi y^3}} \cdot e^{-\frac{1}{2y}}$$

Q3

Change of variable: Weibull distribution

Posted on October 17, 2012 by Jonathan Mattingly | Comments Off

A random variable T has the Weibull(λ, α) if it has probability density function

$$f(t) = \lambda \alpha t^{\alpha-1} e^{-\lambda t^\alpha} \qquad (t > 0)$$

where $\lambda > 0$ and $\alpha > 0$.

1. Show that T^α has an exponential(λ) distribution.
2. Show that if U is a uniform$(0, 1)$ random variable, then

$$T = \left(-\frac{\log(U)}{\lambda}\right)^{\frac{1}{\alpha}}$$

has a Weibull(λ, α) distribution.

$T \sim$ Weibull(λ, α),

$$f(t) = \lambda \alpha t^{\alpha-1} e^{-\lambda t^\alpha}, \quad t>0,\ \lambda>0,\ \alpha>0$$

1) let $Y = T^\alpha$

$$f_Y(y) = \frac{f_T(t)}{\left|\frac{dy}{dt}\right|} = \frac{\lambda \alpha t^{\alpha-1} e^{-\lambda t^\alpha}}{\alpha t^{\alpha-1}}$$

$$\frac{dy}{dt} = \alpha t^{\alpha-1} \qquad = \lambda e^{-\lambda t}$$

$$y = t^\alpha \qquad = \lambda e^{-\lambda}(e^{\frac{1}{\alpha}\ln y})^\alpha$$

$$\ln y = \alpha \ln t$$

$$\Rightarrow \frac{1}{\alpha}\ln y = \ln t \qquad = \lambda e^{-\lambda}(e^{\ln y})$$

$$\Rightarrow t = e^{\frac{1}{\alpha}\ln y} \qquad = \lambda e^{-\lambda y} = \exp(\lambda)$$

2) show if U is uniform $(0,1)$ then

$$T = \left(-\frac{\log(U)}{\lambda}\right)^{\frac{1}{\alpha}}$$

has a weibull distribution.

$$\Rightarrow T^{\alpha} = -\frac{\ln(U)}{\lambda} \Rightarrow \lambda T^{\alpha} = -\ln(U) \Rightarrow U = e^{-\lambda T^{\alpha}}$$

$$f_T(t) = \frac{f_U(u)}{\left|\frac{dt}{du}\right|} = \frac{1}{\frac{du}{dt}} \qquad \frac{du}{dt} = \frac{d}{dt}e^{-\lambda t^{\alpha}}$$

$$f_U(u) = \frac{1}{b-a} = \frac{1}{1-0} = 1$$

$$= \frac{du}{dy}\cdot\frac{dy}{dt} = \frac{d}{dy}\left(e^{-\lambda y}\right)\cdot\frac{d}{dt}\left(t^{\alpha}\right)$$

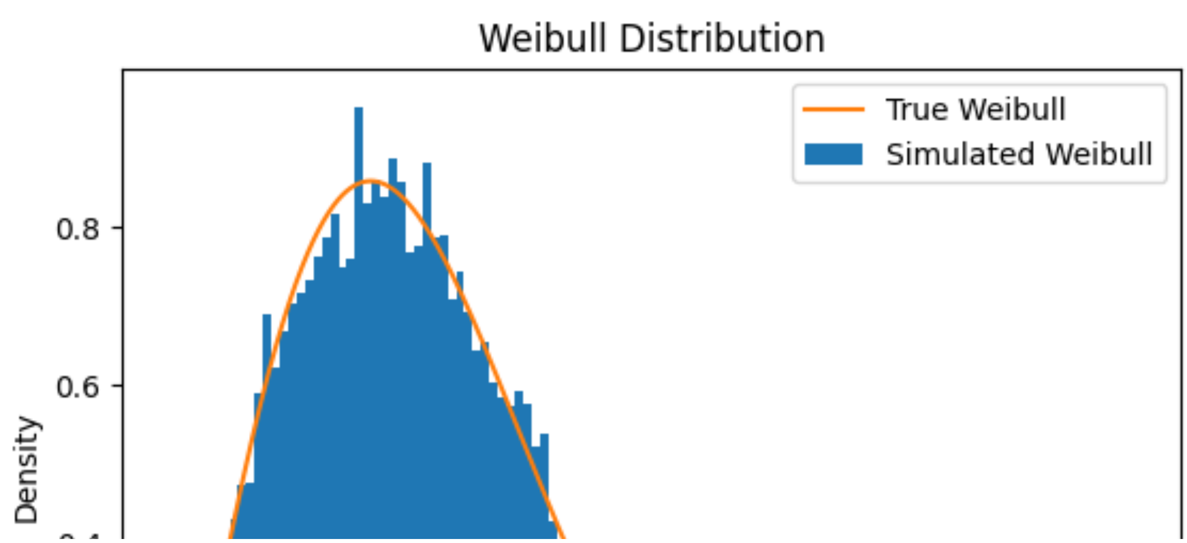
$$y = t^{\alpha} \qquad \frac{dy}{dt} = \alpha t^{\alpha-1}$$

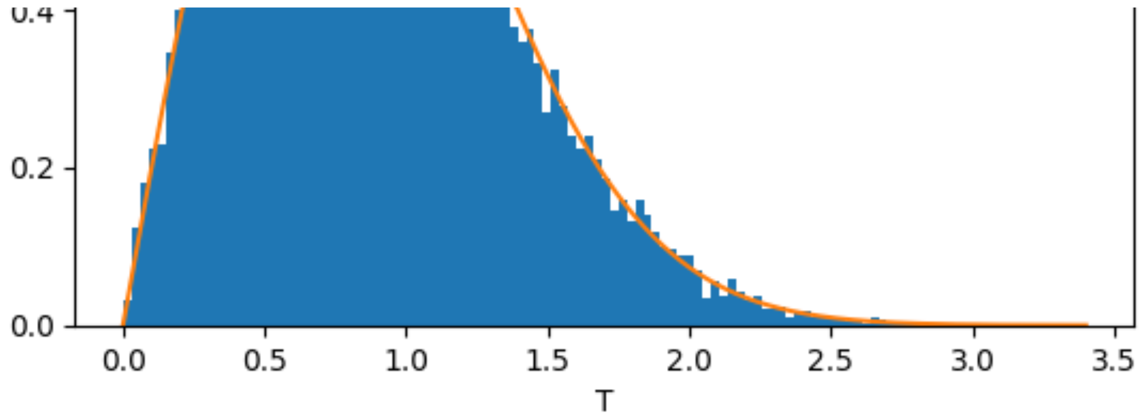
$$= -\lambda e^{-\lambda t^{\alpha}}\cdot\alpha\cdot t^{\alpha-1}$$

$$\therefore f_T(t) = -\lambda e^{-\lambda t^{\alpha}}\cdot\alpha t^{\alpha-1}$$

In [46]:

```python
## simulating a weibull

# samples
num_samples = 15000

# weibull function
def weibull(t, l, alpha):
    return (l) * np.exp(-l * (t ** alpha)) * alpha * (t ** (alpha - 1)

# drawing from uniform distribution
uniform_samples = np.random.uniform(0, 1, num_samples)

# weibull parameters
l = 1
alpha = 2

# T
T = ( ( (-1) * np.log(uniform_samples)) * (1/l) ) ** (1/alpha)

# plotting
plt.hist(T, bins=100, density=True)

#draw a true weibull over it
x = np.linspace(0, 3.4, 15000)
y = weibull(x, l, alpha)
plt.plot(x, y)

#labels
plt.xlabel('T')
plt.ylabel('Density')
plt.title('Weibull Distribution')
plt.legend(['True Weibull', 'Simulated Weibull'])

plt.show()
```





Q4

In [7]:
```
img = Image.open('q4.png')
display(img)

img = Image.open('q4-2.png')
display(img)
```

Exercise 6.2: Monte Carlo Integration from Expected Value II

Return to the integration problems described in Exercise 3.6. Perform the integration using the above scheme based on viewing the integral as an expected value from Section 6.5 using random samples $\mathrm{Uniform}[0, 2]$.

6.5 Monte Carlo Integration : Second Look

The expression in Equation 6.5 provides a different way to approximate an integral that that given in Section 3.3. If we are given a function f whose integral we desire over an interval $[a, b]$ then we can reinterpret the integral as an expected value by introducing the random variable $U \sim \mathrm{Uniform}([a, b])$. Recall that U has a probability density ρ given by

$$\rho(x) = \begin{cases} \frac{1}{b-a} & \text{for } x \in [a, b] \\ 0 & \text{for } x \notin [a, b] \end{cases}$$

Then

$$\int_a^b f(x)dx = (b-a)\int_a^b \frac{f(x)}{b-a}dx = (b-a)\int_{-\infty}^{\infty} f(x)\rho(x)dx = (b-a)\mathbf{E}f(U)$$

The hope would then be that if U_n we a collection of mutually independent random variables each distributed uniformly on the interval $[a, b]$ then

$$\frac{1}{n}\sum_{k=1} f(U_k) \xrightarrow{n\to\infty} \mathbf{E}f(U) = \frac{1}{b-a}\int_a^b f(x)dx$$

We will see that this follows from a generalization of the law of large number as explained in Section 5.6. There only Bernoulli random variables were considered.

Here we have used a uniform random variable. However if ρ is the density of a random variable Y then

$$\int_a^b f(x)dx = \int_a^b \frac{f(x)}{\rho(x)}\rho(x)dx = \int_{-\infty}^{\infty} \frac{f(x)}{\rho(x)}\mathbf{1}_{[a,b]}(x)\rho(x)dx \sim \frac{1}{n}\sum_{k=1} \frac{f(Y_k)}{\rho(Y_k)}\mathbf{1}_{[a,b]}(Y_k)$$

If the Y_k are a collection of mutually independent random variables with probability density ρ.

In [8]:

```
img = Image.open('q4-1.png')
display(img)
```

$$\int_a^b f(x)dx = (b-a)\int_a^b \frac{f(x)}{b-a}dx = (b-a)\int_{-\infty}^{\infty} f(x)\rho(x)dx = (b-a)\mathbf{E}f(U)$$

In [16]:

```
## doing question 4

## trying to estimate integral shown above using above monte carlo exp


#getting samples from uniform distribution

#num samples
sample_size = 100

#integral from a to b
a = 0
b = 2
samples = np.random.uniform(a, b, sample_size)

#function to be integrated
def f(x):
    return (0.5 * (x**3)) - (x ** 2) + 1

# f(U)
f_samples = f(samples)

# Expectation of f(U)
E_f_samples = np.mean(f_samples)

# multiplied by b-a
integral_estimate = (b - a) * E_f_samples

# showing the estimate
print("Estimate of integral is ", integral_estimate)

# actual value of integral
print("Actual value of integral is ", 1.333333)
```

```
Estimate of integral is  1.3164592105913804
Actual value of integral is  1.333333
```

In [23]:

```
## plotting how the estimate changes with number of samples

estimates = []

max_samples = 10000

for i in range(1, max_samples):
    sample_size = i
    a = 0
    b = 2
    samples = np.random.uniform(a, b, sample_size)
    f_samples = f(samples)
    E_f_samples = np.mean(f_samples)
    integral_estimate = (b - a) * E_f_samples
    estimates.append(integral_estimate)

plt.plot(estimates)
plt.xlabel('Number of samples')
plt.ylabel('Estimate of integral')
plt.title('Estimate of integral vs number of samples')
plt.show()
```