ENAE 441 - 0101

HW04: Random Variables

Due on November 11th. 2025 at 09:30 AM

Dr. Martin, 09:30 AM

Vai Srivastava

November 20^{th}, 2025

Problem 1: Standard Deviation

Consider two zero-mean uncorrelated random variables W and V with standard deviations σ_w and σ_v respectively. What is the standard deviation of the random variable $X = W + V$? Note: Do not assume gaussian distribution.

Solution

If W, V are zero-mean and uncorrelated, then

$$\operatorname{Var}(W+V) = \operatorname{Var}(W) + \operatorname{Var}(V) + 2\operatorname{Cov}(W, V) = \sigma_w^2 + \sigma_v^2$$

Hence

$$\boxed{\sigma_X = \sqrt{\sigma_w^2 + \sigma_v^2}}$$

Problem 2: Correlation Coefficient

Consider two scalar RVs X and Y.

(a) Prove that if X and Y are independent, then their correlation coefficient $\rho = 0$.

(b) Find an example of two RVs that are not independent but that have a correlation coefficient of zero.

(c) Prove that if Y is a linear function of X then $\rho = \pm 1$.

Solution

Part A

If X, Y are independent and have finite second moments, then

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

so

$$\boxed{\rho = \frac{\operatorname{Cov}(X, Y)}{\sigma_X \sigma_Y} = 0}$$

(provided $\sigma_X, \sigma_Y > 0$)

Part B

Let $X \sim Unif(-1, 1)$ and $Y = X^2$. Then

$$\mathbb{E}[X] = 0, \quad \mathbb{E}[Y] = \frac{1}{3}, \quad \mathbb{E}[XY] = \mathbb{E}\left[X^3\right] = 0,$$

so $\operatorname{Cov}(X, Y) = 0$ and $\rho = 0$, yet Y is a deterministic function of $X \implies$ not independent.

Part C

If $Y = aX + b$ with $a \neq 0$,

$$\operatorname{Cov}(X, Y) = \operatorname{Cov}(X, aX + b) = a \operatorname{Var}(X), \quad \sigma_Y = |a|\sigma_X$$

so

$$\boxed{\rho = \frac{a \operatorname{Var}(X)}{\sigma_X |a| \sigma_X} = sgn(a) \in \{\pm 1\}}$$

(If $a = 0$, then $\sigma_Y = 0$ and ρ is undefined.)

Problem 3: Probability Density

Consider the following function

$$\mathrm{f_{XY}}\left(x,y\right)=\begin{cases} ae^{-2x}e^{-3y} & x>0, y>0 \\ 0 & \text{otherwise} \end{cases}$$

and answer the following questions.

(a) Find the value of a so that $f_{XY}(x,y)$ is a valid joint probability density function.

(b) Calculate $\bar{x}$ and $\bar{y}$.

(c) Calculate $\mathbb{E}\left[X^2\right], \mathbb{E}\left[Y^2\right]$, and $\mathbb{E}[XY]$.

(d) Calculate the autocorrelation matrix of the random vector $Z=[X,Y]^T$.

(e) Calculate the variance σ_x^2, variance σ_y^2, and the covariance C_{XY}.

(f) Calculate the autocovariance matrix of the random vector $Z=[X,Y]^T$.

(g) Calculate the correlation coefficient between X and Y.

Solution

Part A

$$1=\iint f_{XY}=a\left(\int_0^\infty e^{-2x}dx\right)\left(\int_0^\infty e^{-3y}dy\right)=a\left(\frac{1}{2}\right)\left(\frac{1}{3}\right)=\frac{a}{6}\Rightarrow a=6.$$

Thus $\mathrm{f_{XY}}\left(x,y\right)=6e^{-2x}e^{-3y}$ and factorizes:

$$\boxed{X\sim\ Exp\left(2\right), Y\sim\ Exp\left(3\right)}\qquad\text{independent}$$

Part B

$$\boxed{\bar{x}=\mathbb{E}[X]=\frac{1}{2},\quad \bar{y}=\mathbb{E}[Y]=\frac{1}{3}}$$

Part C

For Exp(λ)$\mathbb{E}\left[X^2\right]=\frac{2}{\lambda^2}$

$$\begin{aligned}\mathbb{E}\left[X^2\right]&=\frac{2}{2^2}=\frac{1}{2}\\ \mathbb{E}\left[Y^2\right]&=\frac{2}{3^2}=\frac{2}{9}\\ \mathbb{E}[XY]&=\mathbb{E}[X]\mathbb{E}[Y]=\frac{1}{2}\cdot\frac{1}{3}\end{aligned}$$

$$\boxed{\mathbb{E}[XY]=\frac{1}{6}}$$

Problem 3 continued on next page...

Part D

$$R_Z = \begin{bmatrix} \mathbb{E}\left[X^2\right] & \mathbb{E}[XY] \\ \mathbb{E}[YX] & \mathbb{E}\left[Y^2\right] \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{6} \\ \frac{1}{6} & \frac{2}{9} \end{bmatrix}$$

Part E

$$\sigma_x^2 = \mathbb{E}\left[X^2\right] - \bar{x}^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$
$$\sigma_y^2 = \frac{2}{9} - \frac{1}{9} = \frac{1}{9}$$
$$C_{XY} = \mathbb{E}[XY] - \bar{x}\bar{y} = \frac{1}{6} - \frac{1}{6} = 0$$

Part F

$$K_Z = \begin{bmatrix} \sigma_x^2 & C_{XY} \\ C_{XY} & \sigma_y^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{9} \end{bmatrix}$$

Part G

$$\rho = \frac{C_{XY}}{\sigma_x \sigma_y} = 0$$

Problem 4: Covariance and Variance

Prove the following two results from lecture where $x \sim \mathcal{N}\left(\bar{x}, \sigma_x^2\right)$ and $e \sim \mathcal{N}\left(0, \sigma_e^2\right)$ and $y = cx + de$.

(a) $\operatorname{Cov}(X, Y) = \mathbb{E}[(x-\bar{x})(y-\bar{y})] = \mathbb{E}[XY] - \bar{x}\bar{y}$

(b) $\operatorname{Var}(Y) = \mathbb{E}\left[(y-\bar{y})^2\right] = c^2\sigma_x^2 + d^2\sigma_e^2$

Solution

Part A

Let $\mu_X = \mathbb{E}[X]$ and $\mu_Y = \mathbb{E}[Y]$.

$$\begin{aligned}
\mathbb{E}\left[(X-\mu_X)(Y-\mu_Y)\right] &= \mathbb{E}\left[XY - \mu_X Y - \mu_Y X + \mu_X\mu_Y\right] \\
&= \mathbb{E}[XY] - \mu_X\mathbb{E}[Y] - \mu_Y\mathbb{E}[X] + \mu_X\mu_Y \\
&= \mathbb{E}[XY] - \mu_X\mu_Y - \mu_Y\mu_X + \mu_X\mu_Y \\
&= \mathbb{E}[XY] - \mu_X\mu_Y
\end{aligned}$$

Thus

$$\begin{aligned}
\operatorname{Cov}(X, Y) = \mathbb{E}\left[(X-\mu_X)(Y-\mu_Y)\right] &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \\
\implies \boxed{\operatorname{Cov}(X, Y) = \mathbb{E}[XY] - \bar{x}\bar{y}}
\end{aligned}$$

Part B

$$\bar{y} = \mathbb{E}[Y] = c\mathbb{E}[x] + d\mathbb{E}[e] = c\bar{x}$$

Then

$$Y - \bar{y} = c(x-\bar{x}) + de$$

Hence

$$\operatorname{Var}(Y) = \mathbb{E}\left[(c(x-\bar{x}) + de)^2\right] = c^2\sigma_x^2 + d^2\sigma_e^2 + 2cd\mathbb{E}[(x-\bar{x})e]$$

If x and e are independent (or just uncorrelated), $\mathbb{E}[(x-\bar{x})e] = 0$, giving

$$\boxed{\operatorname{Var}(Y) = c^2\sigma_x^2 + d^2\sigma_e^2}$$

Problem 5: Central Limit Theorem + Mappings

(a) In python, generate a random variable x_1 distributed by a uniform distribution $x_i \sim \mathcal{U}[-1, 1]$ using `np.random.uniform` function. Sample $N = 10$ points from this distribution, plot those points as a histogram.

(b) Compute the sample mean, $\hat{\mu}$, and sample variance, $\hat{\sigma}^2$, of a sample set using functions `np.mean` and `np.var` functions, and determine if the reported values match the analytic mean and variance, $\mathbb{E}[x_1]$ and $\mathbb{E}\left[(x_1 - \mu_{x_1})^2\right]$ respectively? Repeat using $N = 10^i$ samples where $1 \leq i \leq 6$, reporting the sample mean. Report the values and explain what you observe.

(c) Create three new random variables x_2, x_3, x_4 in addition to x_1, each also distributed from a uniform distribution $\mathcal{U}[-1, 1]$. Sample $N = 100,000$ values from $x_1 - x_4$, and use these independent variables to compute a new set of random variables y_1, y_2, and y_3 defined as

$$y_1 = \frac{x_1 + x_2}{2}$$
$$y_2 = \frac{x_1 + x_2 + x_3}{3}$$
$$y_3 = \frac{x_1 + x_2 + x_3 + x_4}{4}$$

Using the sampled values of x_1 through x_4, plot a histogram $\mathrm{p}(x_1), \mathrm{p}(y_1), \mathrm{p}(y_2), \mathrm{p}(y_3)$.

(d) Explain what you see.

(e) Transform y_3 into two new random variables z and q and plot the resulting distributions of $\mathrm{p}(z)$ and $\mathrm{p}(q)$ next to the original distribution $\mathrm{p}(y_3)$ where

$$z = \mathrm{g}(y) = 2y + 3$$
$$q = \mathrm{f}(y) = e^y$$

(f) Explain how the distribution changes with both transformations, and if it makes sense.

Solution

Part A

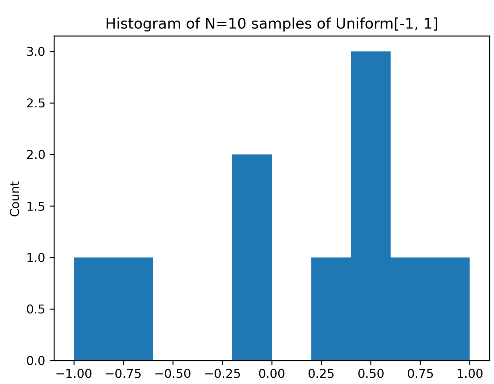


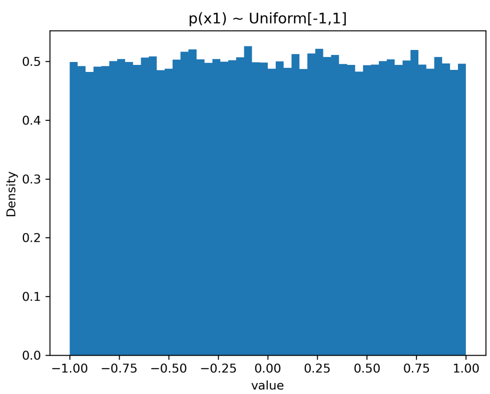
Figure 1: Histogram of $N = 10$ Samples of $x_1 = \mathcal{U}[-1, 1]$

Problem 5 continued on next page...

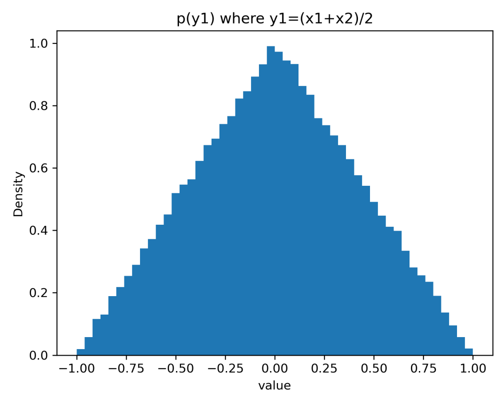
Part B

```
N: [10, 100, 1000, 10000, 100000, 1000000]
Sample Mean: [0.10250805625588852, -0.032626784183775806, 0.0004959945373625319, -0.0024624755192279566, 0.002176523839871754, -0.00031839639155678014]
Sample Variance: [0.2771748942896723, 0.28974256408052684, 0.34257578288265483, 0.33154695575060733, 0.3329566888674737, 0.33320518775561064]
Sample Mean Error: [0.10250805625588852, 0.032626784183775806, 0.0004959945373625319, 0.0024624755192279566, 0.002176523839871754, 0.00031839639155678014]
Sample Variance Error: [0.05615843904366102 4, 0.0435907692528064 7, 0.00924244954932151 9, 0.0017863775827259842, 0.0003766444658596102, 0.00012814557772267143]
The sample mean and variance converge toward 0 and 1/3 as N grows.
The errors shrink similar to O(1/sqrt(N)).
```

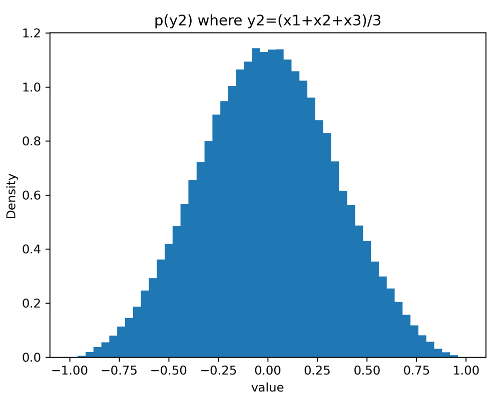
Part C



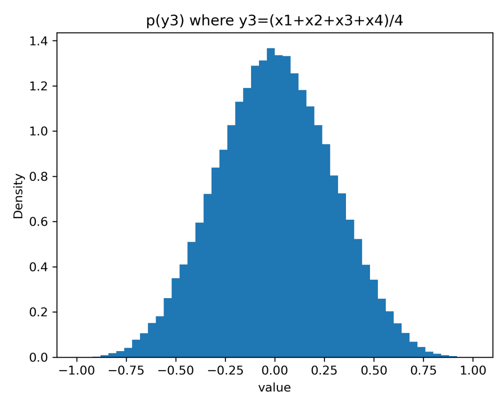
(a) $\mathrm{p}\left(x_1\right)$ where $x_1 = \mathcal{U}[-1, 1]$



(b) $\mathrm{p}\left(y_1\right)$ where $y_1 = \frac{x_1+x_2}{2}$



(c) $\mathrm{p}\left(y_2\right)$ where $y_2 = \frac{x_1+x_2+x_3}{3}$

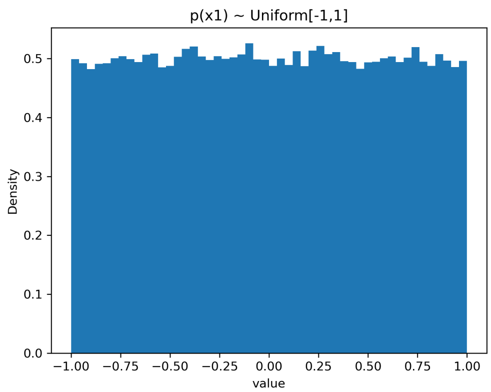


(d) $\mathrm{p}\left(y_3\right)$ where $y_3 = \frac{x_1+x_2+x_3+x_4}{4}$

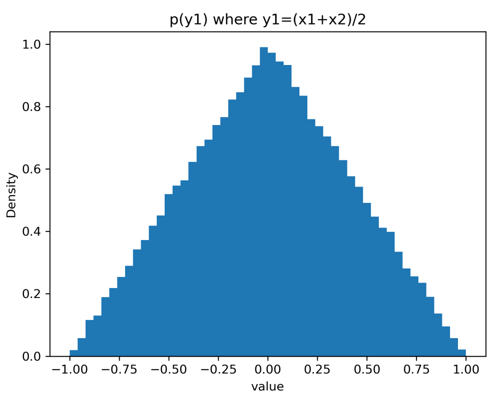
Part D

```
Averages of i.i.d. variables trend toward a normal distribution.
As you average more uniform distributions, the histogram looks more Gaussian and its spread decreases similar to 1/sqrt(N).
```

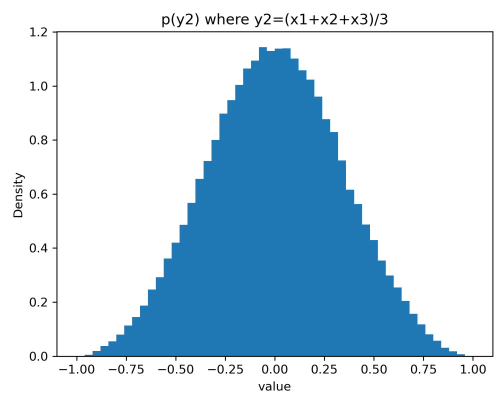
Part E



(a) $\mathrm{p}\left(y_3\right)$ where $y_3 = \frac{x_1+x_2+x_3+x_4}{4}$



(b) $\mathrm{p}\left(z\right)$ where $z = \mathrm{g}\left(y\right) = 2y + 3$



(c) $\mathrm{p}\left(q\right)$ where $q = \mathrm{f}\left(y\right) = e^y$

Part F

```
Linear Transform (z):
    Same shape, shifted right by 3, stretched by factor of 2. Density scales by 1/|2|, so the peak is lower but the curve is simply shifted and stretched.
Non-linear Transform (q):
    Monotone but not shape-preserving. Values within [-1, 1] map to [exp(-1), exp(1)]. The distribution becomes positively skewed, compressing on the left (near 1/e) and stretching on the right (towards e).
```

Code

See the Python code for this assignment.