Chapter 6

Method of Transformations & Order Statistics

Detailed Solutions

6.1 Basic Concept

6.1 Moment Generating Function (Definition)

Problem: $X \sim Exp(1)$, $f(x) = e^{-x}$.
Solution:

(a) **Find MGF:**

$$
\begin{aligned}
M_X(t) = E[e^{tX}] &= \int_0^\infty e^{tx} \cdot e^{-x}\,dx \\
&= \int_0^\infty e^{-(1-t)x}\,dx \\
&= \left[\frac{e^{-(1-t)x}}{-(1-t)}\right]_0^\infty
\end{aligned}
$$

For the integral to converge at ∞, we must have $1 - t > 0 \implies t < 1$.

$$M_X(t) = 0 - \frac{1}{-(1-t)} = \frac{1}{1-t}, \quad t < 1$$

(b) **Calculate Moments:**

$$
\begin{aligned}
M'(t) &= \frac{d}{dt}(1-t)^{-1} = (1-t)^{-2} \\
M''(t) &= \frac{d}{dt}(1-t)^{-2} = 2(1-t)^{-3} \\
E[X] &= M'(0) = (1-0)^{-2} = 1 \\
E[X^2] &= M''(0) = 2(1-0)^{-3} = 2
\end{aligned}
$$

(a) $M_X(t) = \frac{1}{1-t}$ for $t < 1$
(b) $E[X] = 1, E[X^2] = 2$

6.2 Identifying Distributions from MGF

Problem: Identify X.
Solution:

(a) $M_X(t) = (0.3 + 0.7e^t)^{10}$. This matches the form of the Binomial MGF: $(q + pe^t)^n$. Here $n = 10, p = 0.7$. **Result:** $X \sim Binomial(n = 10, p = 0.7)$.

(b) $M_X(t) = e^{5t+8t^2}$. This matches the Normal MGF: $e^{\mu t+\frac{1}{2}\sigma^2 t^2}$. Comparing terms: $\mu = 5$. $\frac{1}{2}\sigma^2 = 8 \implies \sigma^2 = 16$. **Result:** $X \sim Normal(\mu = 5, \sigma^2 = 16)$.

(a) Binomial $(n = 10, p = 0.7)$
(b) Normal $(\mu = 5, \sigma = 4)$

6.3 Jacobian Method (1 RV)

Problem: Formula and Jacobian reasoning.
Solution:

(a) **Formula:**

$$f_Y(y) = f_X(g^{-1}(y)) \cdot \left| \frac{d}{dy} g^{-1}(y) \right|$$

where $x = g^{-1}(y)$.

(b) **Why Absolute Value?** The Probability Density Function (PDF) must always be non-negative ($f(y) \geq 0$). However, if the transformation function $g(X)$ is decreasing (negative slope), the derivative $\frac{dx}{dy}$ will be negative. The absolute value ensures that the resulting probability density remains positive. Conceptually, probabilities are areas, and area cannot be negative.

(a) $f_Y(y) = f_X(x) \left| \frac{dx}{dy} \right|$
(b) To ensure the PDF is non-negative.

6.4 Order Statistics Formulas

Problem: PDFs of Min and Max.
Solution:

(a) **Minimum ($X_{(1)}$):**

$$f_{X_{(1)}}(x) = n[1 - F(x)]^{n-1} f(x)$$

Logic: One variable takes value x (term $f(x)$), and the other $n - 1$ variables must be larger than x (term $[1 - F(x)]^{n-1}$). There are n ways to choose the minimum.

(b) **Maximum ($X_{(n)}$):**

$$f_{X_{(n)}}(x) = n[F(x)]^{n-1}f(x)$$

Logic: One variable takes value x, and the other $n-1$ variables must be smaller than x (term $[F(x)]^{n-1}$).

(a) $n[1-F(x)]^{n-1}f(x)$ (b) $n[F(x)]^{n-1}f(x)$

6.2 Intermediate

6.5 Method of MGF (Sum of Normals)

Problem: Prove $Y = X_1 + X_2$ is Normal.
Solution:

Let $X_i \sim N(\mu_i, \sigma_i^2)$. The MGF is $M_{X_i}(t) = \exp(\mu_i t + \frac{1}{2}\sigma_i^2 t^2)$. Since X_1, X_2 are independent:

$$\begin{aligned} M_Y(t) &= M_{X_1+X_2}(t) = M_{X_1}(t) \cdot M_{X_2}(t) \\ &= \exp\left(\mu_1 t + \frac{1}{2}\sigma_1^2 t^2\right) \cdot \exp\left(\mu_2 t + \frac{1}{2}\sigma_2^2 t^2\right) \\ &= \exp\left((\mu_1 + \mu_2)t + \frac{1}{2}(\sigma_1^2 + \sigma_2^2)t^2\right) \end{aligned}$$

This is the MGF of a Normal distribution with: Mean $= \mu_1 + \mu_2$ and Variance $= \sigma_1^2 + \sigma_2^2$.

Proven.

6.6 Power Transformation

Problem: $f_X(x) = 3x^2$ on $(0, 1)$. Find PDF of $Y = X^3$.
Solution:

1. **Range:** Since $0 < x < 1$, then $0 < x^3 < 1 \implies 0 < y < 1$.
2. **Inverse:** $y = x^3 \implies x = y^{1/3}$.
3. **Jacobian:** $\frac{dx}{dy} = \frac{d}{dy}(y^{1/3}) = \frac{1}{3}y^{-2/3}$.
4. **Substitute:**

$$\begin{aligned} f_Y(y) &= f_X(x)\left|\frac{dx}{dy}\right| \\ &= 3(y^{1/3})^2 \cdot \left|\frac{1}{3}y^{-2/3}\right| \\ &= 3y^{2/3} \cdot \frac{1}{3}y^{-2/3} \\ &= 1, \quad 0 < y < 1 \end{aligned}$$

$f_Y(y) = 1$ for $0 < y < 1$.
$Y \sim Uniform(0, 1)$.

6.7 Inverse Transformation

Problem: $f_X(x) = 2/x^3$ for $x > 1$. Find PDF of $Y = 1/X$.
Solution:

1. **Range:** $x > 1 \implies 0 < 1/x < 1 \implies 0 < y < 1$.
2. **Inverse:** $x = 1/y$.
3. **Jacobian:** $\frac{dx}{dy} = -y^{-2}$. Absolute value is $1/y^2$.
4. **Substitute:**

$$\begin{aligned} f_Y(y) &= \frac{2}{(1/y)^3} \cdot \frac{1}{y^2} \\ &= 2y^3 \cdot \frac{1}{y^2} \\ &= 2y, \quad 0 < y < 1 \end{aligned}$$

This is a Beta(2, 1) or Power function distribution.

$$\boxed{f_Y(y) = 2y, \quad 0 < y < 1.}$$

6.8 Rational Transformation

Problem: $f_X(x) = 2x$ on $(0, 1)$. $Y = \frac{X^2}{1-X^2}$.
Solution:

1. **Range:** As $x \to 0, y \to 0$. As $x \to 1, y \to \infty$. Range is $y > 0$.
2. **Inverse:** $y(1 - x^2) = x^2 \implies y - yx^2 = x^2 \implies y = x^2(1 + y) \implies x^2 = \frac{y}{1+y}$. Since $x > 0$, $x = \sqrt{\frac{y}{1+y}}$.
3. **Jacobian:** Let's use the chain rule with $U = X^2$. Let $U = X^2$. $f_U(u) = f_X(x)|\frac{dx}{du}| = 2\sqrt{u} \cdot \frac{1}{2\sqrt{u}} = 1$ on $(0, 1)$. So $U \sim U(0, 1)$. Now $Y = \frac{U}{1-U}$. Inverse $U = \frac{Y}{1+Y}$. $\frac{du}{dy} = \frac{(1+Y)(1)-Y(1)}{(1+Y)^2} = \frac{1}{(1+Y)^2}$.
4. **PDF of Y:**

$$f_Y(y) = f_U(u) \left| \frac{du}{dy} \right| = 1 \cdot \frac{1}{(1+y)^2}$$

$$\boxed{f_Y(y) = \frac{1}{(1+y)^2}, \quad y > 0.}$$

6.9 System Reliability (Order Statistics)

Problem: n components, $X \sim Exp(\lambda)$.
Solution:

PDF: $f(x) = \lambda e^{-\lambda x}$, CDF: $F(x) = 1 - e^{-\lambda x}$.

(a) **Series System ($Y = \min$):** Using the min formula:

$$\begin{aligned} f_Y(y) &= n[1 - F(y)]^{n-1} f(y) \\ &= n[e^{-\lambda y}]^{n-1}(\lambda e^{-\lambda y}) \\ &= n\lambda e^{-(n-1)\lambda y} e^{-\lambda y} \\ &= (n\lambda) e^{-(n\lambda)y} \end{aligned}$$

This is an Exponential distribution with rate $n\lambda$.

(b) **Parallel System ($Z = \max$):** CDF of max:

$$F_Z(z) = [F(z)]^n = (1 - e^{-\lambda z})^n$$

PDF is the derivative:

$$f_Z(z) = n(1 - e^{-\lambda z})^{n-1}(\lambda e^{-\lambda z})$$

(c) **Comparison: Parallel results in a higher expected lifetime.** Physical Reasoning: In a series system, the weakest link determines the lifetime (if one breaks, all stop). In a parallel system, the system survives as long as at least one component survives (redundancy).

Mathematical Proof: $E[\text{Series}] = \frac{1}{n\lambda}$. $E[\text{Parallel}] = \int_0^\infty (1 - F_Z(z))dz = \int_0^\infty (1 - (1 - e^{-\lambda z})^n)dz$. Alternatively, let $Z = \max$. For exponentials, $E[\max] = \sum_{k=1}^n \frac{1}{k\lambda}$. Since $\sum_{k=1}^n \frac{1}{k} > \frac{1}{n}$, Parallel $>$ Series.

(a) Series: $Exp(n\lambda)$
(b) Parallel CDF: $(1 - e^{-\lambda z})^n$
(c) Parallel $>$ Series (Redundancy increases life).

6.10 Geometric Transformation (Random Chord)

Problem: $X \sim U(0, R)$. Chord L.
Solution:

(a) **Geometry:** Half-chord length is $\sqrt{R^2 - X^2}$. Total length $L = 2\sqrt{R^2 - X^2}$.

(b) **Support:** If $X = 0$ (center), $L = 2R$ (Diameter). If $X = R$ (edge), $L = 0$. Range: $0 < L < 2R$.

(c) **Find PDF:** Inverse: $L/2 = \sqrt{R^2 - X^2} \implies L^2/4 = R^2 - X^2 \implies X = \sqrt{R^2 - L^2/4}$. Jacobian:

$$\frac{dx}{dL} = \frac{1}{2\sqrt{R^2 - L^2/4}} \cdot \left(-\frac{2L}{4}\right) = -\frac{L}{4\sqrt{R^2 - L^2/4}} = -\frac{L}{2\sqrt{4R^2 - L^2}}$$

PDF calculation ($f_X(x) = 1/R$):

$$f_L(l) = \frac{1}{R}\left|-\frac{l}{2\sqrt{4R^2 - l^2}}\right| = \frac{l}{2R\sqrt{4R^2 - l^2}}$$

(d) **Likelihood:** As $l \to 2R$, the denominator $\sqrt{4R^2 - l^2} \to 0$, so $f_L(l) \to \infty$. As $l \to 0$, $f_L(l) \to 0$. It is much more likely to find long chords (near the diameter) than very short ones.

(a) $L = 2\sqrt{R^2 - X^2}$
(c) $f_L(l) = \frac{l}{2R\sqrt{4R^2 - l^2}}$

6.3 Challenge

6.11 Sum of Independent Uniforms (Irwin-Hall)

Problem: $X_1, X_2 \sim U(0, 1)$. PDF of $Z = X_1 + X_2$.
Solution:

This is the convolution of two unit squares.

$$f_Z(z) = \int_{-\infty}^{\infty} f_{X_1}(x) f_{X_2}(z - x) dx$$

The integral is the length of the intersection of line $x_1 + x_2 = z$ with the unit square.

- Case $0 \le z < 1$: $f_Z(z) = \int_0^z 1 \cdot 1 dx = z$.
- Case $1 \le z \le 2$: $f_Z(z) = \int_{z-1}^1 1 \cdot 1 dx = 1 - (z - 1) = 2 - z$.

This forms a triangular distribution.

$$\boxed{f_Z(z) = \begin{cases} z & 0 \le z < 1 \\ 2 - z & 1 \le z \le 2 \\ 0 & \text{else} \end{cases}}$$

6.12 Ratio of Normals (Cauchy)

Problem: $X, Y \sim N(0, 1)$. $V = X/Y$.
Solution:

Using the general formula for ratio $V = X/Y$:

$$f_V(v) = \int_{-\infty}^{\infty} |y| f_{X,Y}(vy, y) dy$$

Since independent, $f_{X,Y} = \frac{1}{2\pi} e^{-(x^2+y^2)/2}$. Substitute $x = vy$:

$$\begin{aligned} f_V(v) &= \int_{-\infty}^{\infty} |y| \frac{1}{2\pi} e^{-(v^2y^2+y^2)/2} dy \\ &= \frac{1}{\pi} \int_0^{\infty} y e^{-\frac{y^2}{2}(1+v^2)} dy \quad \text{(Symmetry)} \end{aligned}$$

Let $u = y^2/2$, $du = ydy$. The exponent is $-(1 + v^2)u$.

$$= \frac{1}{\pi} \int_0^{\infty} e^{-(1+v^2)u} du = \frac{1}{\pi(1 + v^2)}$$

This is the standard Cauchy PDF.

$\boxed{\text{Proven.}}$

6.13 Box-Muller Transformation

Problem: Prove $Z_1, Z_2 \sim N(0,1)$ and independent.
Solution:

Inverse transformation: $U_1 = e^{-(Z_1^2+Z_2^2)/2}$, $U_2 = \frac{1}{2\pi}\arctan(Z_2/Z_1)$. Jacobian $J(z_1, z_2) \to (u_1, u_2)$ involves derivatives of exponential and arctan. Alternatively, calculate Jacobian from $(u_1, u_2) \to (z_1, z_2)$:

$$z_1^2 + z_2^2 = -2\ln u_1 \implies u_1 = e^{-r^2/2}$$

Using polar coordinates logic, the joint PDF transforms to:

$$f_{Z_1,Z_2}(z_1, z_2) = \frac{1}{\sqrt{2\pi}}e^{-z_1^2/2} \cdot \frac{1}{\sqrt{2\pi}}e^{-z_2^2/2}$$

which is the product of two standard normal PDFs.

$\boxed{\text{Proven.}}$

6.14 Range of a Sample

Problem: Range $R = X_{(n)} - X_{(1)}$ for $U(0,1)$.
Solution:

Joint PDF of Min (u) and Max (v) for uniform distribution is:

$$f_{U,V}(u, v) = n(n-1)(v-u)^{n-2}, \quad 0 < u < v < 1$$

Let $R = V - U$ and $M = V$. Then $U = M - R$. Jacobian is 1. Limits: $0 < r < 1$ and $r < m < 1$.

$$\begin{aligned} f_R(r) &= \int_r^1 n(n-1)r^{n-2}\,dm \\ &= n(n-1)r^{n-2}[m]_r^1 \\ &= n(n-1)r^{n-2}(1-r) \end{aligned}$$

This is a Beta distribution $Beta(n-1, 2)$ scaled? No, just the PDF derived.

$\boxed{f_R(r) = n(n-1)(r^{n-2} - r^{n-1}), \quad 0 < r < 1.}$

6.4 Application

6.15 Verifying Transformation via Simulation

Problem: $Y = X^2/(1 - X^2)$.
Solution:

```
import numpy as np
import matplotlib.pyplot as plt

# 1. Generate X ~ 2x
u = np.random.uniform(0, 1, 10000)
x = np.sqrt(u)

# 2. Transform to Y
y = (x**2) / (1 - x**2)

# 3. Plotting
plt.figure(figsize=(10, 6))

# Histogram of Simulation
# Limit range to 0-10 for visibility (PDF has long tail)
plt.hist(y, bins=100, range=(0, 10), density=True,
         alpha=0.6, color='skyblue', label='Simulation')

# Theoretical PDF: f(y) = 1/(1+y)^2
y_vals = np.linspace(0, 10, 1000)
pdf_vals = 1 / (1 + y_vals)**2
plt.plot(y_vals, pdf_vals, 'r-', linewidth=2, label='Theoretical PDF')

plt.xlabel('y')
plt.ylabel('Density')
plt.title('Distribution of Rational Transformation')
plt.legend()
plt.show()
```

6.16 System Reliability Simulation

Problem: Series vs Parallel MTTF.
Solution:

```
import numpy as np
import matplotlib.pyplot as plt

lam = 0.5
n_components = 5
n_sims = 10000

# Generate Data
components = np.random.exponential(scale=1/lam, size=(n_sims,
    n_components))

# Calculate Lifetimes
series_life = np.min(components, axis=1)
```

```python
parallel_life = np.max(components, axis=1)

# Plot
plt.figure(figsize=(10, 6))
plt.hist(series_life, bins=50, alpha=0.5, label='Series (Min)')
plt.hist(parallel_life, bins=50, alpha=0.5, label='Parallel (Max)')
plt.legend()
plt.title('System Lifetime Distribution')
plt.show()

# Mean Time To Failure (MTTF)
print(f"Series MTTF: {np.mean(series_life):.4f} (Theory: {1/(
    n_components*lam)})")
# Theory Parallel: sum(1/(k*lam)) for k=1 to 5
theory_parallel = sum([1/(k*lam) for k in range(1, n_components+1)])
print(f"Parallel MTTF: {np.mean(parallel_life):.4f} (Theory: {
    theory_parallel:.4f})")
```

Result: Parallel system lasts much longer.