3. Derivative factor adjusts for local concentrating and diluting.
   a. See Fig. 24.

   Fig. 24: Transformation from Normal Using Square Root

   y
   x
   Transformation
   X density
   Y density

4. Simplest transformation example
   a. If \( r(x) = cx \) for some constant \( c \)
      i. (as when the new measure is the old measure on a new scale)

5. Argument requires that \( dr^{-1}(y)/dy \) exists on \( Y \).
   a. \( \frac{dr^{-1}}{dy}y^{-1}(y) = 1/r(r^{-1}(y)) \).
      i. By differentiating \( r(r^{-1}(y)) = y \) gives
         \( r'(r^{-1}(y)) \cdot \frac{dr^{-1}}{dy}y^{-1}(y) = 1 \).
   b. A transformation of a continuous variable can have a discrete distribution:
      i. \( X \) is uniform on \((0, 1)\) (i.e., the probability density function is 1 throughout this region) and \( r(x) = 0 \).
      ii. \( Y \) is now discrete, taking only the value \( 0 \),
      iii. the arguments above break down because although the derivative of \( r \) exists, it is zero everywhere, and hence the derivative of \( r^{-1} \) exists nowhere.

6. Argument extends to some cases when \( r \) has a flat spot:
   a. Requires some more care.
   b. Example:
      i. \( X \) uniform on \((0, 1)\) and \( r(x) = (x - .5)^3 \). See Fig. 25.
      ii. Then \( r'(x) = 3(x - .5)^2 \)
      iii. Then \( r^{-1}(y) = y^{1/3} + .5 \) and its derivative is \( \frac{1}{3}y^{-2/3} \).
      iv. The probability density function over the domain \((-5^{1/3}, 5^{1/3})\) is \( 1/(3(3^{1/3} + .5 - .5)^2) = \frac{1}{3}y^{-2/3} \).

   Lecture 11

   Fig. 25: Transformation \((x - 1/2)^3\)

   - The probability density function is defined everywhere on the domain except at zero.
   7. Case with transformation both increasing and decreasing:
      a. Domain \( \mathcal{X} \) of \( X \) splits into disjoint subsets \( \mathcal{X}_j \).
      b. \( r \) monotonic on each of \( \mathcal{X}_j \).
         i. monotonic means either non-increasing, or non-decreasing.
      c. \( f_Y(y) = \sum_{x \in r^{-1}(\{y\})} f_X(x)/r'(x) \).
      d. Example:

   Lecture 11

   107 Lecture 11

   i. \( X \) has probability density function \( f_X(x) \) equalling 1 on \((-1/2, 1/2)\), 0 elsewhere
      • Transformation \( Y = r(X) \) for \( r(x) = x^2 \)
         ▷ non-increasing on \( \mathcal{X}_1 = (-\frac{1}{2}, 0) \), and
         ▷ non-decreasing on \( \mathcal{X}_2 = (0, \frac{1}{2}) \).
      • \( r^{-1}(y) = \sqrt{|y|} \)
      • \( \frac{dr^{-1}}{dy}y^{-1}(y) = \text{sgn}(y) \cdot \frac{1}{2}y^{-1/2} \).
      • \( \text{sgn}(y) = \begin{cases} 1 & \text{if } y > 0 \\ -1 & \text{if } y < 0 \\ 0 & \text{if } y = 0 \end{cases} \)
      • \( f_Y(y) = | -1/(2\sqrt{y}) | + | 1/(2\sqrt{y}) | = y^{-1/2} \) for \( 0 < y \leq 1/4 \), and 0 otherwise.
   ii. Note that formula fails at \( y = 0 \). See Fig. 26.
      • Remember from Riemann integration discussion that value of probability density function at one point doesn’t matter.
   iii. The probability density function diverges to \( \infty \) as \( y \to 0 \).
   iv. Integrals representing probabilities of sets like \((0, b)\) or \([0, b)\) are improper.
   v. Improper integrals are evaluated as limits of well-defined integrals:
      • \( P(Y \leq 1/8) = \int_0^{1/8} y^{-1/2} \, dy \)
      • Integral up to point where probability density function is infinite is taken as limit
        \( \lim_{a \to 0, a > 0} \int_a^{1/8} y^{-1/2} \, dy = \int_a^{1/8} 2y^{1/2} \, dy \bigg|_a^{1/8} = \)
Density

\[ \lim_{a \to 0, a > 0} \sqrt{\frac{1}{2} - 2\sqrt{a}} = \sqrt{\frac{1}{2}}. \]

8. Probability density function may diverge to \( \infty \) in middle.
   a. \( X \) has probability density function \( f_X(x) \) equalling 1 on \((-1/2, 1/2)\), 0 elsewhere
   b. Transformation \( Y = r(X) \) for \( r(x) = \text{sgn}(x)x^2 \)
      i. \( r^{-1}(y) = \text{sgn}(y)\sqrt{|y|} \)
      ii. \( \frac{dy}{dx}r^{-1}(y) = \frac{3}{2}|y|^{-1/2} \)
      iii. \( f_Y(y) = 1/(2\sqrt{|y|}) \) for \(|y| \leq 1/2\), and 0 otherwise. See Fig 27.
      iv. Formula still fails at \( y = 0 \).

Lecture 11

b. Don’t define expectation if \( \int_X |x| f_X(x) \, dx = \infty \).
2. Expectation of transformation of a random variable defined as before
   a. Want \( E(r(X)) \) for some random variable \( X \) taking values in \( X' \).
   b. Transform to new variable \( Y = r(X) \) taking values in \( Y' \).
   c. Calculate its probability density function \( f_Y(y) \)
   d. Report \( E(Y) = \int_{Y'} yf_Y(y) \, dy \).
   e. Can calculate expectation of transformation without constructing new density:
      \( E(r(X)) = \int_X r(x)f_X(x) \, dx \).
      i. As before, \( f_Y(y) = \left| \frac{dy}{dx} \right| f_X(r^{-1}(y)) \)
      ii. \( \int_Y yf_X(r^{-1}(y)) \frac{dy}{dx} \, dy = \int_X r(x)f_X(x) \, dx \)
3. Definition of typical value
   a. Expectation
      i. Advantage: explicitly and uniquely defined.
      ii. Disadvantage: Sometimes isn’t defined.
   b. Median
      i. Advantage: Always defined.
      ii. Disadvantage: Sometimes not unique.
4. Linearity
   a. Let \( Y = aX + b \) for some constants \( a \), \( b \)
   b. Then \( E(Y) = aE(X) + b \).
   i. Use summation and constant multiple rules for integration:

\[ \text{E}(Y) = \int_X (ax + b)f_X(x) \, dx \]
\[ = a \int_X x f_X(x) \, dx + b \int_X f_X(x) \, dx = aE(X) + b \]

5. Other moments as before:
   a. The \( r \)-th moment is defined as \( E(X^r) \).
   b. The \( r \)-th central moment is defined as \( E((X - E(X))^r) \).
6. Describing spread via Variance:
   a. \( V(X) \) is the second central moment: average squared distance from mean.
   b. Alternate formulation:
      \( V(X) = E(X^2) - E(X)^2 \).
   c. Standard deviation: typical distance from expectation:
      \( \text{SD}(X) = \sqrt{V(X)} \)
   d. Linearity:
      \( V(aX + b) = E((aX + b - E(aX + b))^2) \)
      = \( E((aX + b - E(aX) - b)^2) \)
      = \( E(a^2(X - E(X))^2) \)
      = \( a^2V(X) \)
      i. Hence \( \text{SD}(aX + b) = |a| \text{SD}(X) \)
WMS: 4.4
E. Particular Distributions
1. Uniform distribution
   a. In symbols, \( X \sim \text{Unif}(a, b) \).
b. probability density function
\[ f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if } x \in [a, b] \\ 0 & \text{otherwise} \end{cases} \]

c. distribution function
\[ F_X(x) = \begin{cases} 0 & \text{if } x < a \\ \frac{x-a}{b-a} & \text{if } x \in [a, b] \\ 1 & \text{if } x > b \end{cases} \]

i. See Fig. 28.

![Fig. 28: Unif(a, b) Distribution](image)

Vertical scale on two panels is not the same.

d. Expectation \( E(X) = \frac{(a + b)}{2} \).

i. \( E(X) = \int_a^b \frac{x}{b-a} \, dx = \frac{(b^2/2 - a^2/2)}{(b-a)} = \frac{(a+b)}{2} \).

ii. We could have seen this through symmetry.

iii. Median is the same.

e. Variance: \( V(X) = \frac{(b-a)^2}{12} \).

i. \( E(X^2) = \int_a^b \frac{x^2}{b-a} \, dx = \frac{(b^3/3 - a^3/3)}{(b-a)} = \frac{(a^2 + ab + b^2)}{3} \).

ii. \( V(X) = \frac{(a^2 + ab + b^2)}{3} - \frac{(a^2 + 2ab + b^2)}{4} = \frac{(a^2 - 2ab + b^2)}{12} = \frac{(b-a)^2}{12} \).

f. R gives probabilities via \text{punif}, but this is hardly necessary.