5. Can move \( f_X(x) \) at some points without impacting \( F_X(x) \).
   a. Recall probability density function is that function whose integral gives probabilities.
   b. Riemann definition of \( \int_a^b h(x) \, dx \)
      i. Finding sequences of partitions of \([a, b]\)
         - as \([a_0, a_1], \ldots, [a_{k-1}, a_k]\) with \( a_0 = a, a_k = b \)
         - so that \( \sum_{j=1}^k (a_j - a_{j-1}) \max_{x \in [a_{j-1}, a_j]} f \) is minimized,
         - so that \( \sum_{j=1}^k (a_j - a_{j-1}) \min_{x \in [a_{j-1}, a_j]} f \) is maximized.
      ii. Round bracket indicates end point is excluded.
      iii. If upper, lower bounds converge, then common value is the integral. See Fig. 15.
   c. Moving the probability density function at one point induces a change in only one interval.
      i. If this interval shrinks, contribution converges to 0. See Fig. 16.
      d. Hence may ignore values of probability density function at finite number of points.
   d. Hence we will not distinguish between integrals over ranges like \([a, b], (a, b), (a, b), [a, b]\).
      i. Square bracket indicates end point is included.
      ii. Round bracket indicates end point is excluded.

B. Quantile Function as a description of the distribution.
1. Quantile Definition:
   a. Heuristically, inverse of \( F_X(x) \).
      i. See Fig. 17.
   b. If unique inverse exists, it is this inverse. See Fig. 18.
      i. Denote the quantile \( p \) of random variable \( X \) by \( \phi_p = F_X^{-1}(p) \).
   c. (At least one) solution to \( F_X(\phi_p) = p \) exists if \( F_X \) continuous by intermediate value theorem.
   d. Inverse is unique if \( f_X(x) > 0 \) except at isolated points.
      i. Ambiguous if \( F_X(x) \) has a flat spot so that various \( x \) satisfy \( F_X(x) = p \). See Fig. 19.
      ii. Then every \( x \) value in flat part satisfies quantile function. See Fig. 20.

5. Can move \( f_X(x) \) at some points without impacting \( F_X(x) \).
   a. Recall probability density function is that function whose integral gives probabilities.
   b. Riemann definition of \( \int_a^b h(x) \, dx \)
      i. Finding sequences of partitions of \([a, b]\)
         - as \([a_0, a_1], \ldots, [a_{k-1}, a_k]\) with \( a_0 = a, a_k = b \)
         - so that \( \sum_{j=1}^k (a_j - a_{j-1}) \max_{x \in [a_{j-1}, a_j]} f \) is minimized,
      ii. Round bracket indicates end point is excluded.
      iii. If upper, lower bounds converge, then common value is the integral. See Fig. 15.
   c. Moving the probability density function at one point induces a change in only one interval.
      i. If this interval shrinks, contribution converges to 0. See Fig. 16.
      d. Hence may ignore values of probability density function at finite number of points.
   d. Hence we will not distinguish between integrals over ranges like \([a, b], (a, b), (a, b), [a, b]\).
      i. Square bracket indicates end point is included.
      ii. Round bracket indicates end point is excluded.

B. Quantile Function as a description of the distribution.
1. Quantile Definition:
   a. Heuristically, inverse of \( F_X(x) \).
      i. See Fig. 17.
   b. If unique inverse exists, it is this inverse. See Fig. 18.
      i. Denote the quantile \( p \) of random variable \( X \) by \( \phi_p = F_X^{-1}(p) \).
   c. (At least one) solution to \( F_X(\phi_p) = p \) exists if \( F_X \) continuous by intermediate value theorem.
   d. Inverse is unique if \( f_X(x) > 0 \) except at isolated points.
      i. Ambiguous if \( F_X(x) \) has a flat spot so that various \( x \) satisfy \( F_X(x) = p \). See Fig. 19.
      ii. Then every \( x \) value in flat part satisfies quantile function. See Fig. 20.

iii. Use \( \phi_p = \inf \{ x \mid F_X(x) \geq p \} \).
iv. Discrete cases might not have a solution to \( p = F_X(\phi) \). See Fig. 21.
   e. Since inverse need not exist, \( \phi_p \) chosen so that \( P(X \leq \phi_p) \geq p, P(X > \phi_p) \geq (1 - p) \).

2. Some important quantiles
   a. The median is the 1/2 quantile.
      i. Measures the center of a distribution.
      ii. The median corresponding to a random variable \( X \) with distribution function \( F_X \) is that value \( \nu_X \) such that \( P(X \leq \nu_X) \geq .5 \) and \( P(X \geq \nu_X) \geq .5 \).
   b. A quartile is a .25 or .75 quantile.
      i. Distinguished by calling upper or lower.

3. Examples of quantiles:
   a. Binomial variables with the probability function
1. Univariate Transformations via the Definition
   a. $X$ takes values in some set $\mathcal{X}$
   b. $r$ is a function defined on $\mathcal{X}$
   c. Want to describe distribution of $Y = r(X)$.
      i. $Y$ takes values in $\mathcal{Y} = r(\mathcal{X})$.
      ii. via the probability function or probability density function for $Y$.
      iii. When $X$ is discrete, this was easy.

b. Then the median is
   \[
   \begin{cases} 
   0 & \text{if } \pi < .5 \\
   1 & \text{if } \pi > .5 \\
   \text{anything} & \text{otherwise.}
   \end{cases}
   \]
   i. Binomial variable, probability function $p(x; 5, 5)$:
      See Fig. 21.

4. Comparison of median and expectation
   a. Disadvantages relative to expectation:
      i. The median can’t be given explicitly, but only as the
         solution to an equation involving bounds on integrals or sums,
      ii. sometimes isn’t unique,
      iii. sometimes doesn’t give much information.
   b. Advantage:
      i. always exists.
      
      WMS: 6.3a

C. Transformations of random variables.

\[ p(x; \pi, 1) = \begin{cases} 
1 & \text{if } x = 1 \\
\pi^x (1 - \pi)^{1-x} & \text{otherwise}
\end{cases} \]

\[ F(x) = \begin{cases} 
0 & \text{if } x \leq \text{probability} \\
\text{anything} & \text{otherwise.}
\end{cases} \]

- Sum over appropriate values of $X$:
  \[ p_Y(y) = P(Y = y) = P(r(X) = y) = \sum_{r(x) = y} p_X(x). \]

- If $r$ is one to one,
  \begin{itemize}
  \item an inverse for $r$ exists on $\mathcal{Y}$,
  \item sums above all have only one addend.
  \end{itemize}

- If $r$ is non-decreasing, then expression in terms of the distribution function is easy:
  \[ F_Y(y) = P(Y \leq y) = P(r(X) \leq y) \]
Lecture 10

1. Probability density function of a transformed continuous variable

a. When $X$ is continuous then generally (but not always) $Y$ also has a probability density function.
   - Let $f_X$ be the probability density function for $X$.
   - Denote the transformation function by $r(x)$.

b. $f_Y(y)$ can generally be expressed in terms of the original probability density function:
   - $f_Y(y) = f_X(r^{-1}(y)) \left| \frac{dy}{dx} \right| r^{-1}(y)$.
   - When $r$ is non-decreasing and has inverse,
     - Differentiate distribution function:
     \[ f_Y(y) = \frac{d}{dy} F_X(r^{-1}(y)) \]
     \[ = F_X(r^{-1}(y)) \frac{d}{dy} r^{-1}(y) \]
     \[ = f_X(r^{-1}(y)) \frac{d}{dy} r^{-1}(y) \] (2)
   - Requires transformation to have positive derivative.

c. Examples:
   - $r(x) = x^2$, probability density function for $X$ is $\exp(-x)$ for $x \geq 0$.

Fig. 23: Transformation of Normal

\[ f_Y(y) = \frac{d}{dy} (1 - F_X(r^{-1}(y))) \]
\[ = -F_X'(r^{-1}(y)) \frac{d}{dy} r^{-1}(y) \]
\[ = f_X(r^{-1}(y)) \left( - \frac{d}{dy} r^{-1}(y) \right) \]
\[ = f_X(r^{-1}(y)) \left| \frac{dy}{dx} \right| r^{-1}(y) \] (4)

2. Probability density function of a transformed continuous variable

- $r^{-1}(y) = \sqrt{y}$, probability density function for $X$ is $c \exp(-x^2)$ for $x \geq 0$.
- $c = 2/\sqrt{\pi}$, but we don’t need this.
- $r^{-1}(y) = y^2$, $\frac{dy}{dx} r^{-1}(y) = 2y$.
- $f_Y(y) = 2c \exp(-y^4)$. See Fig. 23.

- Can express using the integration change of variables formula:

Fig. 22: Transformation of Exponential

- $r(x) = \sqrt{x}$, probability density function for $X$ is stretched over a wide range.
- Probability density function of $Y$ should be lower than if $r$ were moving more slowly.

f. Absolute values around derivative account for the case if $r$ non-increasing instead of non-decreasing:
   - By the definition of the distribution function,
     \[ F_Y(y) = P(Y \leq y) = P(r(X) \leq y) \]
     \[ = P(X \geq r^{-1}(y)) = 1 - F_X(r^{-1}(y)). \] (3)
   - Probability density function becomes
     \[ f_Y(y) = \frac{d}{dy} (1 - F_X(r^{-1}(y))) \]
     \[ = -F_X'(r^{-1}(y)) \frac{d}{dy} r^{-1}(y) \]
     \[ = f_X(r^{-1}(y)) \left( - \frac{d}{dy} r^{-1}(y) \right) \]
     \[ = f_X(r^{-1}(y)) \left| \frac{dy}{dx} \right| r^{-1}(y) \] (4)

i. $\int_A f_X \, dx = \int_B f_X \, \frac{dx}{dy} \, dy$;
ii. as functions of $y$ using $r$ then the product $f_X \frac{dx}{dy}$ satisfies requirements for a probability density function.

e. Interpretation: probability density function of $Y$ at $y$ has two factors:
   - if it inherits form the distribution for $X$,
   - part that arises because of stretching or contracting the scale.
   - If $r$ is moving very quickly as $x$ moves, then probability arising from $f_X$ is stretched over a wide range.

- $f_Y(y) = \exp(-\sqrt{y}) \frac{1}{2\sqrt{y}}$. See Fig. 22.