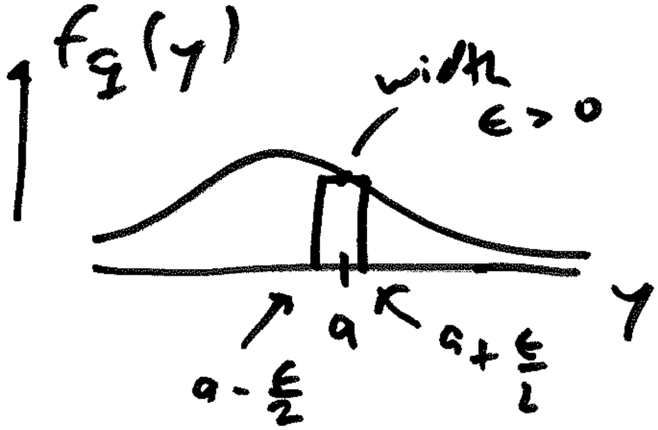
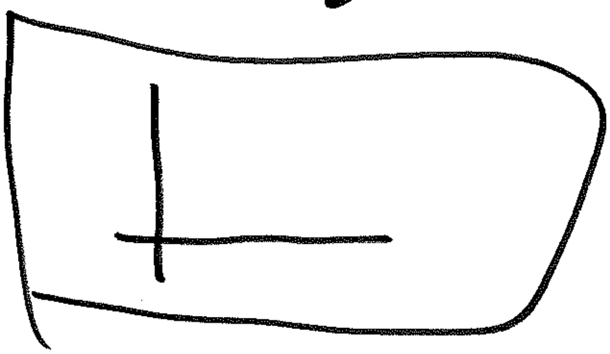
This time: PDFs, CDFs,
next time: inverse CDFs

AMS 131
9 Aug 19

(1)

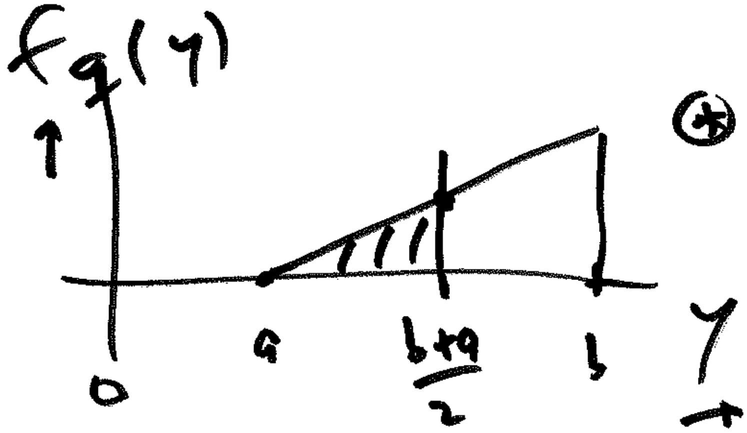


$$P\left(a - \frac{\epsilon}{2} \le Y \le a + \frac{\epsilon}{2}\right) = \int_{a - \frac{\epsilon}{2}}^{a + \frac{\epsilon}{2}} f_Y(y)\, dy$$



$$= \epsilon \cdot f_Y(a)$$

so
$$f_Y(a) = \frac{P\left(a - \frac{\epsilon}{2} \le Y \le a + \frac{\epsilon}{2}\right)}{\epsilon}$$



$$\circledast = P\left(a \le Y \le \frac{b+a}{2}\right)$$

$$= \int_a^{\frac{b+a}{2}} \frac{2(y-a)}{(b-a)^2}\, dy$$

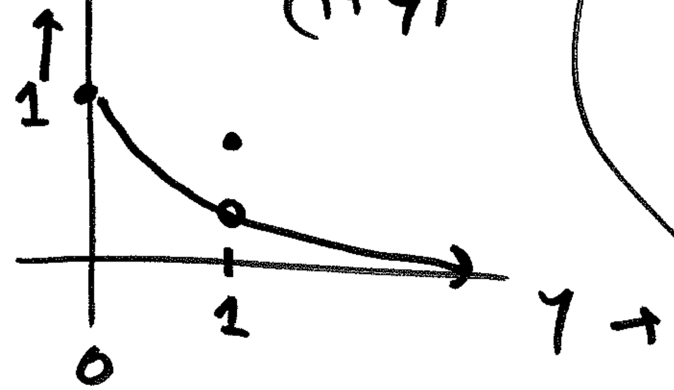
$$\circledast = \frac{1}{2}\left(\frac{b+a}{2} - a\right) h$$

$$h = \frac{2\left(\frac{b+a}{2} - a\right)}{(b-a)^2}$$

$$= \frac{1}{4}$$

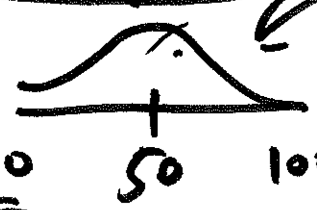
shapes of probability distributions ②

$$f(y) = \frac{1}{(1+y)^2}$$



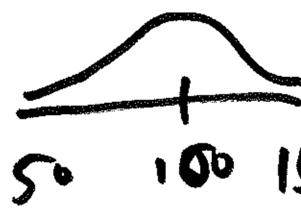
same spread

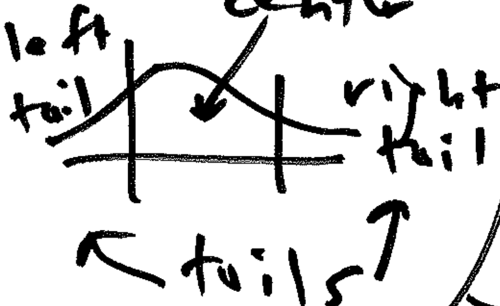
PDF sketch



same shape

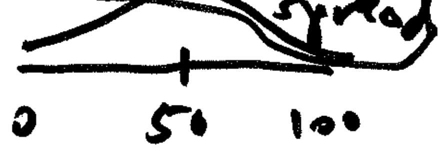
different center



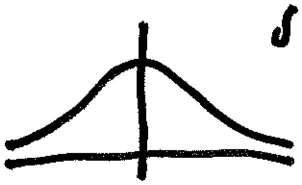


same shape

different spread



symmetric



same center

~~point of~~ symmetry

right skewed

(long right-hand tail)

(positive skew)

asymmetric = skewed

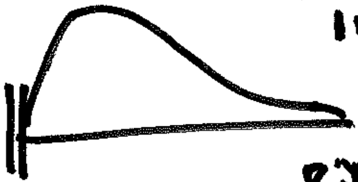
left skewed

long left-hand tail (negative skew)

same center

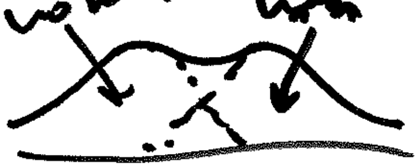
same spread

different shape



family income (yearly)

PDF of scores on THT1 (in %)



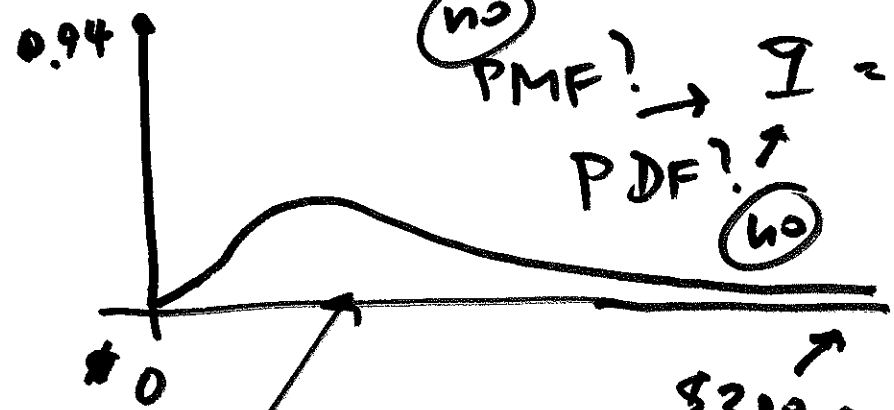
height of 131 students

(9.55)

mixture distribution across gender

gender ← latent variable

latent variable ← unobserved

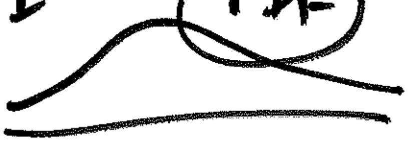


(no) PMF? → Y = total GMB ← $ of merchandise bought in (say) 4-week period

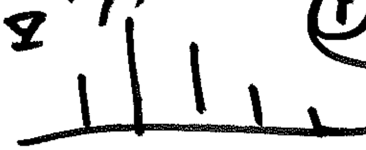
PDF? (no)

(mixed) dist: partly discrete, partly continuous

$f_Y(y)$ (PDF) continuous RV



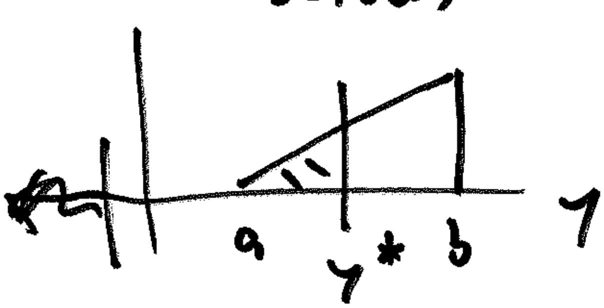
$f_Y(y)$ (P)MF discrete RV



spike plot

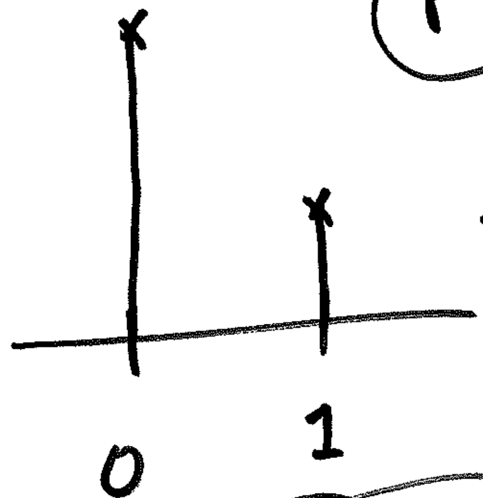
(CDF) = cumulative (probability) distribution function (at or below)

$F_Y(y) = P(Y \le y)$



$$F_Y(y) = \begin{cases} 0 & \text{for } y \le a \\ \int_a^y f_Y(y)\,dy & a \le y < b \\ 1 & \text{for } y \ge b \end{cases}$$

$p < \frac{1}{2}$

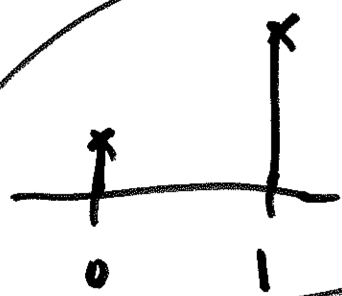


$(Y \mid p) \sim \text{Bernoulli}(p)$

PMF, discrete

$$f_Y(y) = \begin{cases} p & \text{if } y = 1 \\ 1-p & \text{if } y = 0 \\ 0 & \text{else} \end{cases}$$

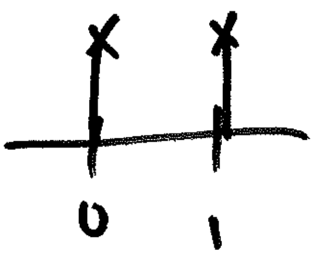
$p > \frac{1}{2}$



$$= p^y (1-p)^{1-y} I_{\{0,1\}}(y)$$

CDF of Y:

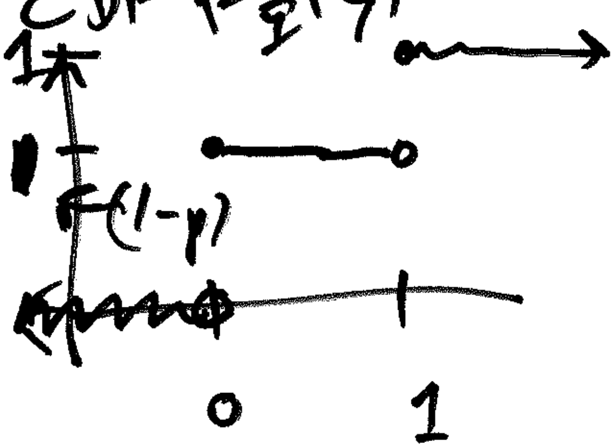
$p = \frac{1}{2}$



$$F_Y(y) = P(Y \le y)$$

$$= \begin{cases} 0 & \text{for } y < 0 \\ (1-p) & 0 \le y < 1 \\ 1 & y \ge 1 \end{cases}$$

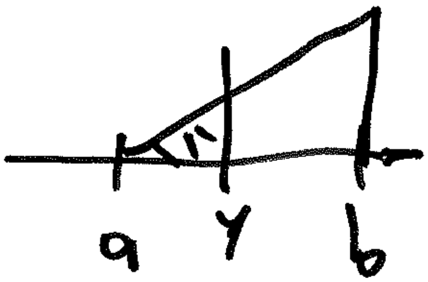
CDF $F_Y(y)$



step function — all discrete RV have step-fn CDFs

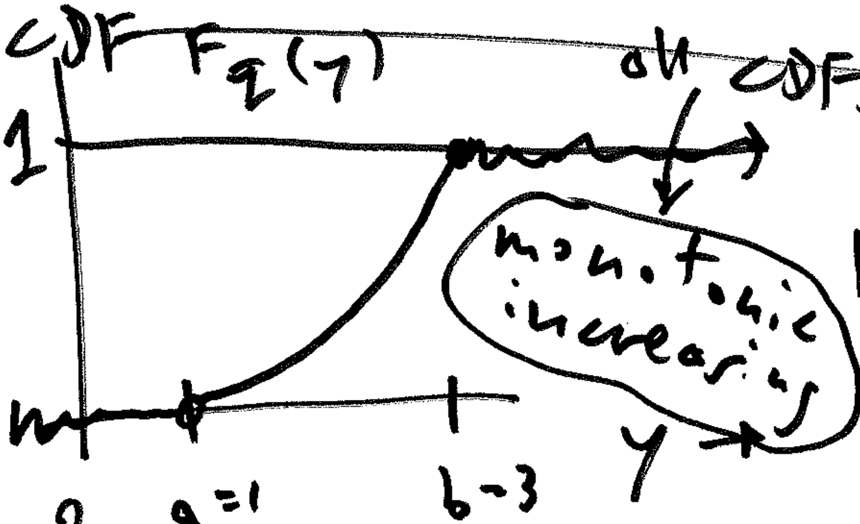
F_Y nondecreasing

$(Y \mid a, b) \sim$ (continuous) TriangularUp(a, b) $\quad (a < b)$



← PDF

$$f_Y(y) = \begin{cases} \dfrac{2(y-a)}{(b-a)^2} & \text{for } a \le y \le b \\ 0 & \text{else} \end{cases}$$



all CDFs for continuous RV

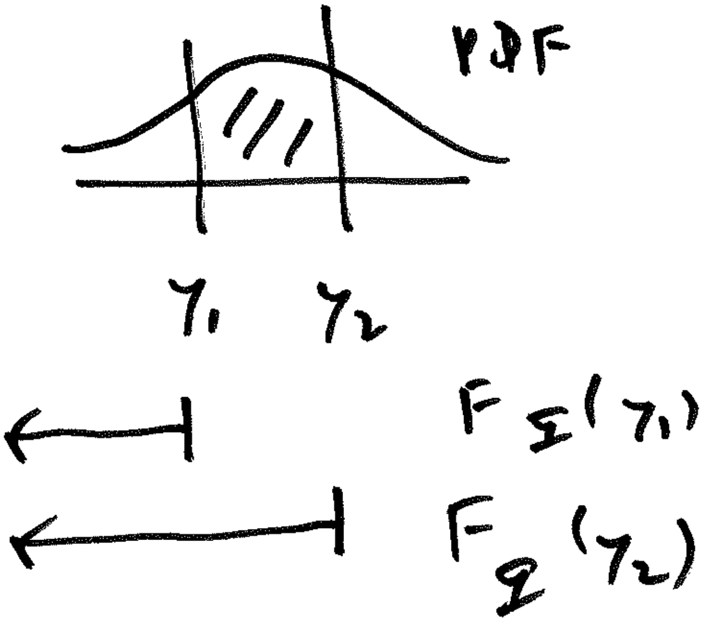
$$F_Y(y) = \begin{cases} 0 & \text{for } y \le a \\ \boxed{?} & a \le y \le b \\ 1 & y \ge b \end{cases}$$

for $a \le y \le b$

$$F_Y(y) = P(Y \le y) = \int_a^y \frac{2(t-a)}{(b-a)^2}\, dt$$

$$= \left(\frac{y-a}{b-a} \right)^2$$

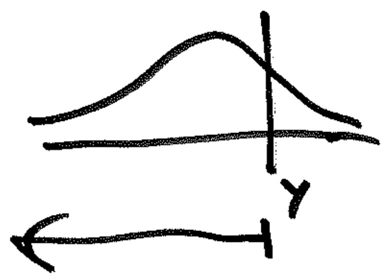
$$P(y_1 < Y \le y_2) = F_Y(y_2) - F_Y(y_1)$$



(10.52)

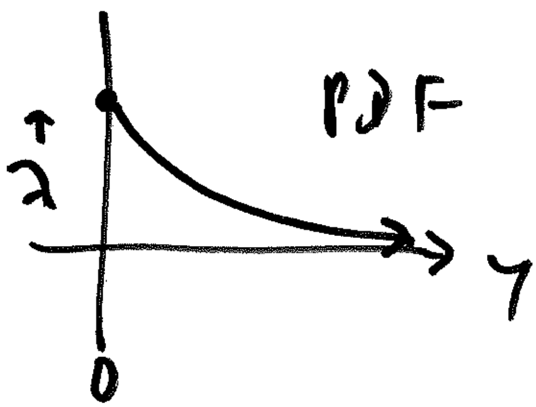
$$F_Y(y) = P(Y \le y)$$

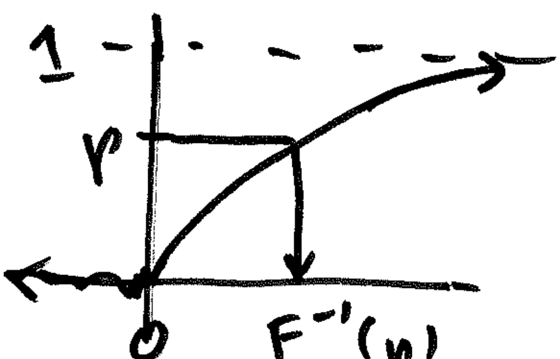
$$P(Y > y) = 1 - F_Y(y)$$

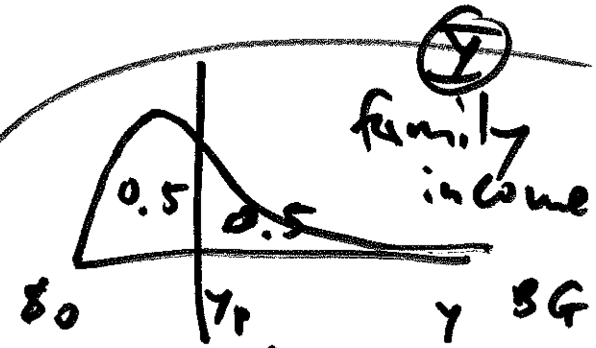


$(\lambda > 0)$

$(Y \mid \lambda) \sim \text{Exponential}(\lambda)$



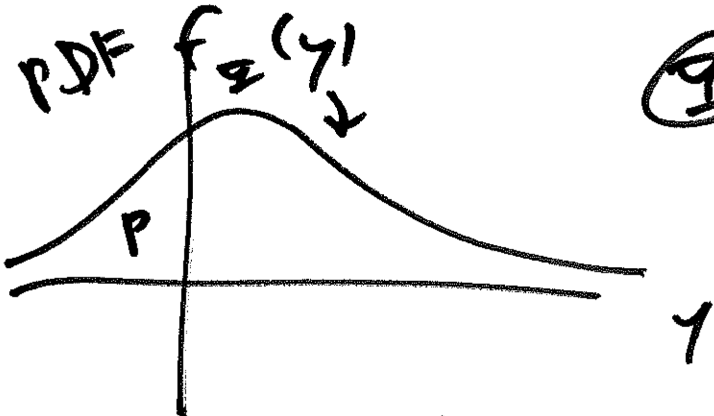


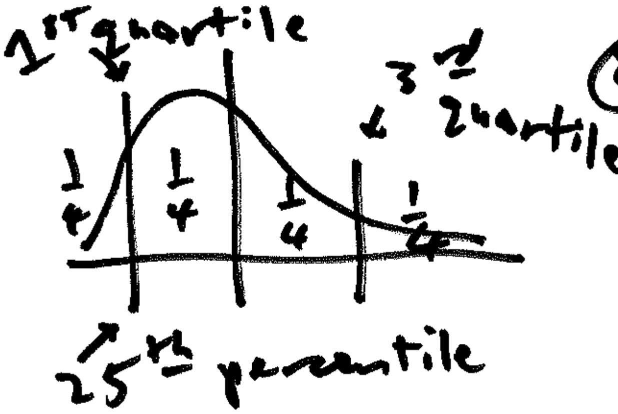


median (50/50 point)

$$F_Y(y_p) = \frac{1}{2}$$

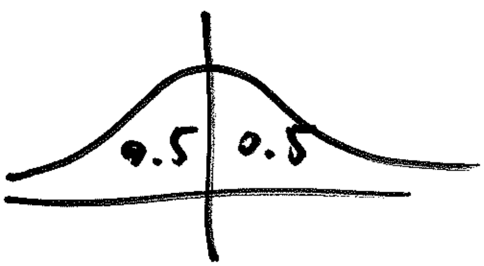
$$\text{median} = y_p = F_Y^{-1}\left(\frac{1}{2}\right)$$

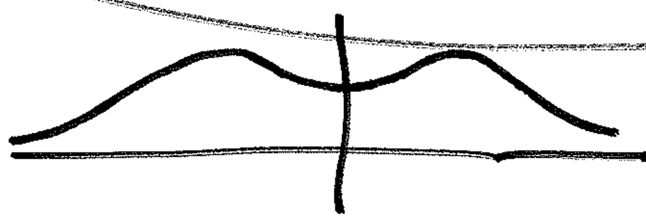




Y_p = p^{th} quantile of dist. of Y

= $100p^{th}$ percentile of





center of symmetry = median

same ✓