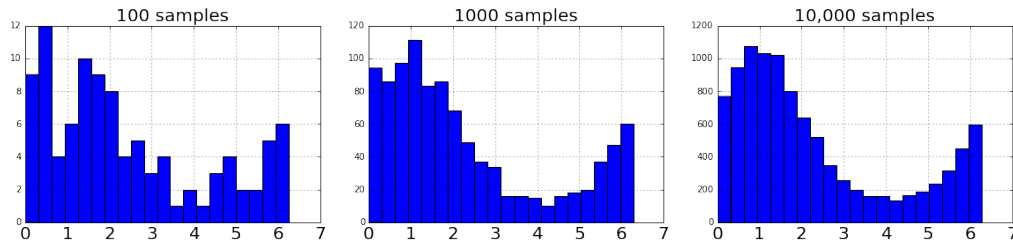
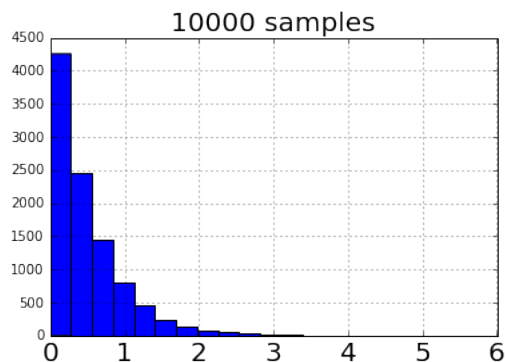


You are allowed one 8 1/2" x 11" sheet of paper with hand-written notes on both sides.

- 1) Below are 3 histograms. Each histogram is generated by sampling from the same unknown distribution a different number of times. Draw a sketch of the unknown distribution.



- 2) For the data in the histogram below, is the mean or median larger? Why?



- 3) Assume the dataframe `nba` contains information about all NBA players for the 2012-2013 season.

player	position	age	team	games	games_started	free_throws	pts
Quincy Acy	SF	23	TOT	63	0	35	171
Steven Adams	C	20	OKC	81	20	79	265
Jeff Adrien	PF	27	TOT	53	12	76	362
Arron Affalo	SG	28	ORL	73	73	274	1330
Alexis Ajinca	C	25	NOP	56	30	56	328
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
Tyler Zeller	C	24	CLE	70	9	87	399

- (a) Write a piece of code to compute the mean age of all players.
- (b) The column `pts` contains the number of points scored by each player during the season. Write a piece of code to compute the standard deviation of the total points scored.
- (c) Write a piece of code that creates a new dataframe called `player_sample` containing 30 randomly selected players from the `nba` dataframe.