Class Notes 2-22-2019

- **The Dartboard Again:**
  What is the expected distance R form the center under uniform impact points?
  We had $f_R(t) = 2t/r_0^2 \, I_{[0,r_0)}$, thus

$$E(R) = \int_{-\infty}^{\infty} t f_R(t) dt = \int_0^{r_0} t \frac{2t}{r_0^2} dt = \frac{2}{3} \left. \frac{t^3}{r_0^2} \right|_0^{r_0} = \frac{2}{3} r_0$$

  not so intuitive, but somewhat.

- **Infinite and Nonexistent Expectations:**

  – **St. Petersburg Paradox (Nicolas Bernoulli (1687-1759)):**
  You repeatedly flip a fair coin until you get your first `H`. If you win on the first flip you get \$2, if you win on the second flip you get double, i.e.,\$4 $= 2^2$, then stop. If you get your first `H` on the n^{th} flip you get \$$2^n$. If X is the amount won when you stop, what is $E(X)$?

$$E(X) = \sum_{n=1}^{\infty} 2^n \frac{1}{2^n} = 1 + 1 + 1 + \ldots = \infty \quad \text{but} \quad P(X < \infty) = 1$$

  – **Continuous Example with Infinite Expectation:**
  Let $f_X(x) = x^{-2} I_{[1,\infty)}(x)$ then

$$E(X) = \int_{-\infty}^{\infty} t f_X(x) dx = \int_1^{\infty} x \frac{1}{x^2} dx = \lim_{b\to\infty} \int_1^b \frac{1}{x} dx = \lim_{b\to\infty} \ln(b) = \infty$$

  – **Nonexistent Expectation:**
  You and I take turns flipping a fair coin until we see the first `H`. If that happens (first `H`) when n is odd I get \$$2^n$ from you. When it happens when n is even, I pay you \$$2^n$. With X denoting my winnings we have $P(X = 2^n) = 2^{-n}$ when n is odd and $P(X = -2^n) = 2^{-n}$ when n is even. Then $P(|X| < \infty) = 1$ and

$$E(X) = 2^1 2^{-1} - 2^2 2^{-2} + 2^3 2^{-3} - 2^4 2^{-4} + \ldots = 1 - 1 + 1 - 1 + \ldots = \text{undefined}$$

- **Expectation of a Function of a Random Variable:**
  Taking a function g of a random variable X creates another random variable $g(X)$. It still maps any $\omega \in \Omega$ into $\mathbb{R}$.

- **Example: Fair Die Roll**
  Suppose we roll a fair die with X denoting the outcome $\in \{1, 2, \ldots, 6\}$. Suppose we win or lose the amount W, namely $W = -1$ if $X = 1, 2, 3$, $W = 1$ if $X = 4$ and $W = 3$ is $X = 5, 6$. Thus we have $W = g(X)$ with $g(1) = g(2) = g(3) = -1$, $g(4) = 1$ and $g(5) = g(6) = 3$. Compute $E[g(X)]$ in two equivalent ways, the first finds the distribution of W and the second uses

$$E[g(X)] = \sum_{x=1}^{6} g(x) P(X = x)$$

  This second method works in general and saves the effort of finding the pmf or pdf for $g(X)$.

Let g be a real valued function defined over the range of a random variable X. If X is a discrete random variable with pmf $f_X(k)$ then

$$E[g(X)] = \sum_k g(k)P(X = k) = \sum_k g(k)f_X(k)$$

If X is a continuous random variable with pdf $f_X(x)$ then

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f_X(x)dx$$

The proof in the discrete case is like the one in the above example, with more general notation. The proof in the continuous case is beyond this course but is plausible if we approximate the continuous random variable by a discretized version (like rounding to some decimals).

- **Expected Randomly Broken Stick Length:**
  A stick of length ℓ is broken at a uniformly random chosen point along $[0, \ell]$. Find the expected length of the longer piece. If the stick is broken at location $X \in [0, \ell]$ then the length of the longer piece is $g(X) = \ell - X$ when $0 \leq X < \ell/2$ and $g(X) = X$ when $X \in [\ell/2, \ell]$. Thus

$$E[g(X)] = \int_0^{\ell} g(x)f(x)dx = \int_0^{\ell/2} \frac{\ell - x}{\ell}dx + \int_{\ell/2}^{\ell} \frac{x}{\ell}dx = \frac{3\ell}{4}$$

- n^{th} **Moment** $E(X^n)$**:**

in the discrete case with pmf $f_X(x)$ we define $\quad E(X^n) = \sum_k k^n f_X(k) = \sum_k k^n P(X = k)$

in the continuous case with pdf $f_X(x)$ we define $\quad E(X^n) = \int_{-\infty}^{\infty} x^n f_X(x)dx$

The second moment $E(X^2)$ is also called the mean square.

- **Uniform Moments:**
  Let $U \sim \text{Unif}[0, c]$ with density $f(u) = \frac{1}{c}I_{[0,c]}(u)$. Then

$$E(U^n) = \int_0^c u^n \frac{1}{c}du = \frac{u^{n+1}}{c(n+1)}\bigg|_0^c = \frac{c^n}{n+1}$$

- **Insurance Example Revisited:**
  This example was neither entirely continuous nor discrete, it was a mix of the two. We have no definition for the expected value of X, the amount that Peter has to pay. However, we can finesse this. Recall that the damage in an accident is $Y \sim \text{Unif}[100, 2000]$ and Peter has to pay Y when $Y \leq 500$ and just 500 when $Y > 500$, thus $X = \min(Y, 500) = g(Y)$, where Y is a continuous random variable and we have definition for $E[g(Y)]$. Thus

$$\begin{aligned} E(X) &= E[g(Y)] = \int_{100}^{2000} \frac{1}{1900}g(y)dy = \int_{100}^{500} \frac{1}{1900}ydy + \int_{500}^{2000} \frac{500}{1900}dy \\ &= \frac{1}{1900}\frac{y^2}{2}\bigg|_{100}^{500} + \frac{500 \cdot 1500}{1900} = \frac{2400}{38} + \frac{7500}{19} = \frac{17400}{38} = 457.89 \end{aligned}$$

- **The Median of a Distribution:**
  An alternate indicator for the center of a distribution is the **median**, defined as any value m such that $P(X \geq m) \geq 1/2$ and $P(X \leq m) \geq 1/2$. m may not be unique.
  It is a better indicator of of the center when X has extremely large values ($\pm$).

- **Example with Extreme Outlier Value:**
  Let X have possible values $\{-100, 1, 2, \ldots, 9\}$ each having equal probability 1/10. Then $E(X) = -100/10 + 1/10 + 2/10 + \ldots + 9/10 = -5.5$. The -100 value pulls the mean way below the 0.9 probability sitting on $1, \ldots, 9$. However, the median can be any any number $m \in [4, 5]$, which gives a more appropriate evaluation for the center of the distribution.
  The median is often used when dealing with incomes as a fairer summary assessment of a population.

- **p-Quantiles:**
  The pth quantile x_p of a random variable X is any number x_p which satisfies
  $P(X \leq x_p) \geq p$ and $P(X \geq x_p) \geq 1 - p$. x_p may not be unique. For $p = .25$ and $p = .75$ the values of x_p are also called first and third quartiles. Deciles and percentiles cut it even finer.

- **The Variance of a Random Variable:**
  The variance captures the amount of variation of X around the mean μ, presumed to be finite. It is defined as

$$\sigma^2 = \mathrm{var}(X) = E[(X - \mu)^2], \qquad \sigma = \sqrt{\mathrm{var}(X)} = \textit{standard deviation} \text{ of } X$$

  The variance is easier to deal with mathematically than the more intuitive variation measure *mean absolute deviation* $E[|X - \mu|]$.
  In a way, σ is meant to counter the distortion of the square function used in the variance.

- **Variance Computation and Alternate Forms:**
  In the discrete case with pmf $p_X(k)$

$$\begin{aligned}
\mathrm{var}(X) &= \sum_k (k-\mu)^2 P(X = k) = \sum_k (k-\mu)^2 p_X(k) = \sum_k k^2 p_X(k) - \sum_k 2k\mu p_X(k) + \sum_k \mu^2 p_X(k) \\
&= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2 = E(X^2) - [E(X)]^2 = \text{mean square} - \text{squared mean}
\end{aligned}$$

  In the continuous case with pdf $f_X(x)$

$$\begin{aligned}
\mathrm{var}(X) &= \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)dx = \int_{-\infty}^{\infty} x^2 f_X(x)dx - \int_{-\infty}^{\infty} 2x\mu f_X(x)dx + \int_{-\infty}^{\infty} \mu^2 f_X(x)dx \\
&= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2 = E(X^2) - [E(X)]^2 = \text{mean square} - \text{squared mean}
\end{aligned}$$

  The alternate expression $E(X^2) - [E(X)]^2$ for $\mathrm{var}(X)$ holds generally.
  $\mathrm{var}(X) = 0 \Leftrightarrow P(X = c) = 1$ for some c, in which case $E(X) = c$.