HW#3 Due next Tuesday
p. 119 #3, #7
p. 150 #1, #2, #4
p. 173 #2, #3
p. 182 #3, #4
p. 238 #1, #3
p. 254 #2
p. 327 #9
p. 339 #2, #6

\(K\)th moment \(E[X^K], k=1, 2, \ldots\)

If all moments exist, then they characterize the random variable. But they need not exist. The integral or sum may diverge.

**Example**

\[
f(x) = \begin{cases} 
0 & , x < 1 \\
\frac{1}{x^2} & , x \geq 1 
\end{cases}
\]

How would we check that it is a density?

- non-negative

\[
\int_{-\infty}^{\infty} f(x) \, dx = 1, \quad \therefore \int_{-\infty}^{\infty} f(x) \, dx = \int_{1}^{\infty} \frac{dx}{x^2} = - \frac{1}{x} \bigg|_{x=1}^{x=\infty} = - \lim_{x \to \infty} \frac{1}{x} + 1 = +1 \; \text{is a valid density.}
\]
Compute $E[X] = \int_{-\infty}^{+\infty} x f(x) \, dx$

$\ln x \Big|_{x=1}^{x=\infty}$ diverges for $x \to \infty$

Pareto distributions useful for computer engineering
Internet traffic has $\infty$ variance $\Rightarrow$ doesn't have a 2nd moment

Height can be thought of as a random variable
so we have a distribution
Weight is RV

Look at two distributions jointly - not necessarily independent

Jointly distributed random variables $X, Y, Z$

$F(x, y) \overset{\text{def}}{=} P\{X \leq x, Y \leq y\}$ for all $x, y \in \mathbb{R}$.

$5'10" \quad 190 \text{ lbs}$

If both $X$ and $Y$ are discrete
then consider joint prob. mass function

$PMF = p(x, y) = P\{X = x, Y = y\}$ joint distribution of heights & weights
If RV's are jointly continuous, then define joint PDF

\[
F(u, v) = \int_{-\infty}^{u} \int_{-\infty}^{v} f(x, y) \, dx \, dy
\]

for all \( u, v \in \mathbb{R} \)

Assume \( p(x, y) \) \( \leftrightarrow \) table of probabilities

\( X \) being jointly height & weight

Then \( F(u, v) = \sum_{y < v} \sum_{x \leq u} p(x, y) \).

Again, not necessarily independent.

Looking at "marginal distribution" of heights - initially you were given more information (joint distribution of heights & weights) and you only want distribution of weights.

Suppose given \( p(x, y) \), want \( p_x(x) = \sum_{y} p(x, y) \).

\[
p_y(y) = \sum_{\text{all } x} p(x, y)
\]

Indicates the marginal distribution w.r.t. \( x \)

In the continuous domain, the marginal distributions

\[
f_x(x) = \int_{-\infty}^{\infty} f(x, y) \, dy \quad \text{and} \quad f_y(y) = \int_{-\infty}^{\infty} f(x, y) \, dx.
\]
Consider: 
F(x, y) CDF

How to get marginal?

\[ F_x(x) = \lim_{y \to \infty} F_{x,y}(x, y) \]

\[ F_y(y) = \lim_{x \to \infty} F_{x,y}(x, y) \]

\[ \{ (1,1,0,0), (1,0,1,0), (0,1,1,0), (1,0,0,1), (0,1,0,1), (0,0,1,1) \} \]

\[ X = \# \text{ of polls until we find the 1st ready} \]
\[ Y = \# \text{ of polls until we find the 2nd ready} \]

<table>
<thead>
<tr>
<th>x</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>( p_x )</th>
</tr>
</thead>
</table>
| 1 | \( \frac{1}{6} \) | \( \frac{1}{6} \) | \( \frac{1}{6} \) | \( \frac{1}{2} \) | \\
| 2 | 0 | \( \frac{1}{6} \) | \( \frac{1}{6} \) | \( \frac{1}{3} \) |
| 3 | 0 | 0 | \( \frac{1}{6} \) | \( \frac{1}{6} \) |

\[ p_{Y \mid x} = \left( \begin{array}{c}
\frac{1}{6} \\
\frac{1}{3} \\
\frac{1}{2}
\end{array} \right) \]

\[ \text{Marginal for } Y \text{ alone} \]

What is \( F(2, 3) = \text{Prob that I have 1, 2 and...} \)

\[ p(1,2) + p(1,3) + p(2,2) + p(2,3) \]
New:

Assume

\( X, Y \) jointly distributed,

\( q(x, y) \) function of R.V.s \( X, Y \).

\[
E[ q(x, y) ] = \begin{cases} 
\sum_{(x,y)} q(x,y) p(x,y) & \text{discrete} \\
\int_{\infty}^{\infty} \int_{\infty}^{\infty} q(x,y) f(x,y) \, dx \, dy & \text{jointly-continuous}
\end{cases}
\]

\( X, Y \) are independent if any of the following hold:

- \( F(x,y) = F_X(x) F_Y(y) \) for all real values \( x \) and \( y \).

- If joint discrete, \( p(x, y) = p_x(x) p_y(y), \forall x, y \in \mathbb{R} \).

- If jointly continuous

\[
f(x,y) = f_X(x) f_Y(y), \forall x, y \in \mathbb{R}.
\]

Example of many joint R.V.s

- \( N_1 \) # of processes using driver 1 (device 1) (Jackson's theorem)
- \( N_2 \) # of processes using driver 2
- \( N_3 \) open network
- Evaluate performance: \( P(n_1,n_2,n_3, \ldots) \)
  \[
  = \prod_{i} P(n_i).
  \]
Look at disk I/O:

- N = total # of users in system
- Nq = how many are in queue
- W_s = service time
- W_q = time I sit in the queue
- W = total time in system (total I/O)

Given \( p_N(n) = p^N q^N \) for \( n = 0, 1, 2, \ldots \)

Compute \( E[N] = \sum_{n=0}^{\infty} n p_N(n) \) expected number of users of this drive

\[
\text{Var}[N] = E[(N - E[N])^2]
\]

\[
= \sum_{n=0}^{\infty} (n - E[N])^2 p_N(n)
\]

\[
= \sum_{n=0}^{\infty} (n^2 - 2nE[N] + E[N]^2) p_N(n)
\]

\[
= \sum_{n=0}^{\infty} n^2 p_N(n) - 2 E[N] \sum_{n=0}^{\infty} n p_N(n) + (E[N])^2 \sum_{n=0}^{\infty} p_N(n)
\]

\[
= \frac{\text{2nd moment of } N}{E[N]} - E[N] \frac{\text{mean of } N}{E[N]} + (E[N])^2 \frac{1}{E[N]}
\]

This is a number
\[
\text{Var} [N] = E[N^2] - (E[N])^2, \quad \text{know this relationship.}
\]

**Example**

Imagine \( W \) total I/O time \( t \) \( \text{min} \succ t_{I/O} \succ t_{\text{max}} \)

\[
W = \int f_w(t) dt
\]

Compute mean I/O time so compute \( E[W] \)

\[
E[W] = \int_{-\infty}^{+\infty} t f_w(t) dt
\]

\[
= \int_{t_{\text{min}}}^{t_{\text{max}}} t f_w(t) dt
\]

What is variance of I/O time?

\[
\text{Var}[W] = E[(W - E[W])^2]
\]

\[
= \int_{-\infty}^{+\infty} \left( t - E[W] \right)^2 f_w(t) dt
\]

\[
= \int_{t_{\text{min}}}^{t_{\text{max}}} \left( t^2 - 2t E[W] + (E[W])^2 \right) f_w(t) dt
\]

\[
= \int_{t_{\text{min}}}^{t_{\text{max}}} t^2 f_w(t) dt - 2E[W] \int_{t_{\text{min}}}^{t_{\text{max}}} t f_w(t) dt + (E[W])^2 \int_{t_{\text{min}}}^{t_{\text{max}}} f_w(t) dt
\]

\[
\]
\[ \text{Var}[W] = E[W^2] - (E[W])^2 \]

new.

Bean counting: charge people for use of compute resources
associate cost \( C(W) \), for I/O time

Assume cost is linear: \( C(W) = cW, \ c > 0 \).

Compute mean cost:
\[
E[C(W)] = \int_{-\infty}^{t_{\text{max}}} C(t) f_W(t) \, dt
\]
\[
= \int_{t_{\text{min}}}^{t_{\text{max}}} c t f_W(t) \, dt
\]
\[
= c E[W].
\]

Variance of cost:
\[
\text{Var}[C(W)] = \int_{-\infty}^{+\infty} \left( c t - c E[W] \right)^2 f_W(t) \, dt
\]
\[
= \int_{t_{\text{min}}}^{t_{\text{max}}} \left( c^2 t^2 - 2 c t E[W] + c^2 (E[W])^2 \right) f_W(t) \, dt
\]
\[
= c^2 E[W^2] - c^2 (E[W])^2
\]
\[
= c^2 \text{Var}[W].
\]