Random variables $X: \mathcal{N} \rightarrow \mathbb{R}$

Distributions (joint, marginal, conditional)

Expectation $E[X] = \sum_{x} x P\{X = x\}$

Linearity, method of indicators

Interpretations

tail-sum formula
means for geometric, negative binomial
Markov inequality $E g(X)$, $E g(X, Y)$
Multi. rule ($X + Y$ independent)

Simple (Math.) example of scaling

<table>
<thead>
<tr>
<th>values $x$</th>
<th>0</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P{X = x}$</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
</tr>
</tbody>
</table>

| $EX = 1$ |
| $\text{var} X = E(X - 1)^2 = (0 - 1)^2 \cdot \frac{1}{2} + (2 - 1)^2 \cdot \frac{1}{2} = 1$ |
| $SD X = \sqrt{1} = 1$ |
| $\sqrt{3} X$ |

<table>
<thead>
<tr>
<th>values $y$</th>
<th>0</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P{Y = y}$</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
</tr>
</tbody>
</table>

| $EY = 3$ |
| $\text{var} Y = E(Y - 3)^2 = E(9) = 9$ |
| $SD(Y) = 3$ |

Note: $\text{var} X = \sum_{x} (x - \mu)^2 P\{X = x\} \geq 0$

Q: Is $\text{var} X = 0$ possible? How?

Recall $\text{var} X = E(X^2) - (EX)^2$.

Thus, $E(X^2) = (EX)^2$

if and only if

100 Exam scores
mean = 60
$SD = 10 = \sqrt{\sum_{i=1}^{100} (x_i - 60)^2 / 100}$

At most _____ scores $< 50$ or $> 70$.
At most _____ scores $< 40$ or $> 80$.
At most _____ scores $< 30$ or $> 90$.

What if score dist. approximately normal?
Additional Rule for Variances

If \( X_1, X_2, \ldots, X_n \) are independent, then
\[
\text{var}(X_1 + \cdots + X_n) = \text{var}(X_1) + \cdots + \text{var}(X_n)
\]

\[
\text{IF } X \text{ and } Y \text{ are independent, then }
\text{var}(X + Y) = \text{var}(X) + \text{var}(Y)
\]

\[
P\left\{ X \geq \mu + \frac{1}{k} \right\} = P\left\{ Y \geq \mu + \frac{1}{k} \right\}
\]

\[
S \text{ of Add. Rule p199}
\]

\[
S = X + Y
\]

\[
\text{E}[S] = \text{E}[X] + \text{E}[Y]
\]

\[
\text{var}[S] = \text{var}[X] + \text{var}[Y]
\]

\[
\text{Square Root 2-way}
\]

\[
\text{E}[X_1 + \cdots + X_n] = n \mu
\]

\[
\text{Var}[X_1 + \cdots + X_n] = n \sigma^2
\]

\[
\text{Special case:}
\]

\[
\text{Binomial}(n,p)
\]

\[
\text{Sn} = X_1 + \cdots + X_n
\]

\[
\text{Sn} = X_1 + \cdots + X_n
\]

\[
X_1 = \frac{X_1}{n}
\]

\[
\frac{X_1}{n}
\]

\[
\frac{X_1}{n}
\]

\[
\frac{X_1}{n}
\]

\[
\frac{X_1}{n}
\]

\[
\frac{X_1}{n}
\]

\[
\frac{X_1}{n}
\]

\[
\frac{X_1}{n}
\]

\[
\frac{X_1}{n}
\]

\[
\frac{X_1}{n}
\]

\[
\frac{X_1}{n}
\]

\[
\frac{X_1}{n}
\]
Law of Averages (p.195)
If n large, \( \bar{X}_n \) probably close to \( \mu \).

\[ \text{PS: } \frac{SD(\bar{X}_n)}{\sqrt{n}} = \frac{\sigma}{\sqrt{n}} \rightarrow 0 \text{ as } n \uparrow \]

By Chebyshev, a r.v. is unlikely to be more than a couple of SD's from its mean.

Central Limit Theorem (+) p.194
If \( X_1, X_2, \ldots, X_n \) independent, same dist., mean \( \mu \), SD \( \sigma \) < \( \infty \), then for large \( n \),

\[ P\{a \leq \frac{S_n - n\mu}{\sigma \sqrt{n}} \leq b\} \approx \Phi(b) - \Phi(a) \]

for any numbers \( a \leq b \).

If \( S_n \) has (consecutive) integer values, then approx. better with continuity correction.

Remember:
\[ E(X+Y) = E(X) + E(Y) \]
\[ E(aX) = a E(X) \]
\[ \text{var}(X+Y) = \text{var}(X) + \text{var}(Y) \] for \( X \) and \( Y \) independent
\[ \text{var}(aX) = a^2 \text{var}(X) \]

Note:
Standardized \( S_n = \) Standardized \( \bar{X}_n \)

\[ \frac{S_n - n\mu}{\sqrt{n} \sigma} = \frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}} \]

(Divide numerator and denominator of LHS by \( n \) to get RHS.)
Ex. 100 die tosses \( X_1, X_2, \ldots, X_{100} \)
\[ \mu = \mathbb{E} X_1 = 3.5 \]
\[ \text{var} \ X_1 = \mathbb{E}(X_1^2) - (\mathbb{E} X_1)^2 = \frac{1^2 + 2^2 + \ldots + 6^2}{6} - (3.5)^2 \]
\[ S_{100} = X_1 + \ldots + X_{100}, \quad \overline{X} = \frac{S_{100}}{100} = \frac{35}{12} \]
\[ \mathbb{E} S_{100} = n\mu = 350 \]
\[ \text{var} S_{100} = n\sigma^2 = \frac{3500}{12} = 291.7 \]
\[ \text{SD}(S_{100}) = \sqrt{n\sigma} = 17.1 \]
\[ \mathbb{E} \overline{X} = 3.5, \quad \text{var} \overline{X} = \frac{\sigma}{\sqrt{n}} = 0.2917 \]
\[ \text{SD} \overline{X} = \frac{\sigma}{\sqrt{n}} = 0.71 \]

Recall
\[ \mathbb{E} g(X) = \sum_x g(x) P\{X = x\} \]
so for \( X = \text{die roll} \),
\[ \mathbb{E} X^2 = \sum_{x=1}^6 x^2 P\{X = x\} = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + \ldots + 6^2 \cdot \frac{1}{6} = \frac{91}{6} \]

Q1. \( P\{320 \leq S_{100} \leq 370\} \)
\[ \approx P\{329.5 < S_{100} < 370.5\} = P\{\frac{329.5 - 350}{17.1} < \frac{S_{100} - 350}{17.1} < \frac{370.5 - 350}{17.1}\} \]
\[ \approx P\{-1.20 < Z < 1.20\} = 0.7698 \]
Q2. \( P\{3.3 \leq \overline{X} \leq 3.7\} \)
\[ \approx \]
\[ (\text{w/o cont. corr.} 0.7584) \]
Ex. For $W$, weight of a random man, 18-24, $E(W) = 162$, $\sigma = 30$ pounds.

If you get a random sample of 25 such men, what is $P\left\{ \frac{\bar{W}}{25} > 170 \right\}$?

$= P\left\{ \frac{\bar{W} - 162}{\frac{30}{\sqrt{25}}} > \frac{170 - 162}{\frac{30}{\sqrt{25}}} \right\}$ (standardize and get same thing.

$n$ big enough here.

$\approx P\left\{ Z > \frac{8}{6} = 1.33 \right\}$

$= .0912$ Note: Cont. Corr.

NOT appropriate!!!

Remark. More general CLT

If $X_1, \ldots, X_n$ (approx.) independent, and every $X_k - \mu_k$ negligible relative to $SD(S_n) = \sqrt{\frac{n}{1} \sigma_k^2}$, then $S_n$ is approx. normal, (*perhaps diff. dist.'s)

mean $\frac{\sum_{k=1}^{n} \mu_k}{\sqrt{\frac{n}{1} \sigma_k^2}}$

Post - Section 3.4 Wrinkles:

Infinite sum rule:

If events $A_1, A_2, A_3, \ldots$ disjoint, then $P(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k)$

Moments: $EX = \sum_{x} x \cdot P\{X = x\}$ provided $\sum_{x} |x| P\{X = x\} < \infty$

(provided at least one)

(See Chapter 4, especially page 268.)
St. Petersburg Paradox
Start with $\$1$.
- Flip coin until first Heads, then stop.
- Every coin flip doubles your money!
- $X$ = # dollars you end up with.

values: 2, 4, 8, 16, 32, ...
probabilities: $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, ...$

$E[X] = \sum_x x P[X = x] = 1 + 1 + 1 + 1 + ... = \infty.$

Variation: Toss die, then coin as above.
- Die odd: Buffet pays Gates $X$ dollars.
- Die even: Gates pays Buffet $X$ dollars.

$W$ = Gates' net winnings = $\pm X$
values of $W$: $-4, -2, +2, +4, ...$
probabilities: $\frac{1}{8}, \frac{1}{4}, \frac{1}{4}, \frac{1}{8}, ...$

$E[W] = \sum_w w P[W = w]$

$= \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} + ...$

$= ?$ Undefined.