1. Hypergeometric distribution (you know this already, just not by name)

2. Conditional expectation/ Law of total expectation (not on test) MIT Book: Sections 18.4.5-18.4.6,

3. Joint distributions of random variables (not on test)
[BT] Section 2.5

20 types of questions possible on midterm. You know how to answer 12. Professor picks 10 at random.

X: number of correct answers

Pr(get all 10 right) ?

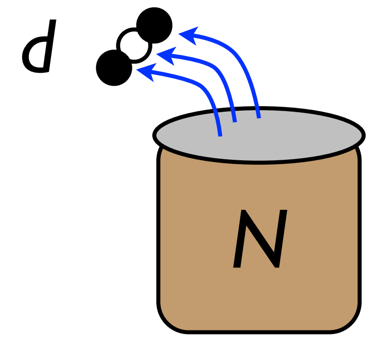
Pr(get 8 of 10) ?

E(number of correct answers) ?

balls in urns – the hypergeometric distribution

B&T, exercise 1.61

Draw *d* balls (without replacement) from an urn containing *N*, of which *w* are white, the rest black.



Let X = number of white balls drawn

$$P(X=i) = \frac{\binom{w}{i}\binom{N-w}{d-i}}{\binom{N}{d}}, \quad i = 0, 1, \ldots, d$$

(note: n choose k = 0 if k < 0 or k > n)

E[X] = dp, where p = w/N (the fraction of white balls)

proof: Let X_j be 0/1 indicator for j-th ball is white, $X = \Sigma X_j$

The X_j are *dependent*, but $E[X] = E[\Sigma X_j] = \Sigma E[X_j] = dp$

Var[X] = dp(1-p)(1-(d-1)/(N-1))

Supreme Court case: Berghuis v. Smith

If a group is underrepresented in a jury pool, how do you tell?

Justice Breyer [Stanford Alum] opened the questioning by invoking the binomial theorem. He hypothesized a scenario involving "**an urn with a thousand balls, and sixty are red, and nine hundred forty are black, and then you select them at random… twelve at a time.**" According to Justice Breyer and the binomial theorem, if the red balls were black jurors then "**you would expect… something like a third to a half of juries would have at least one black person**" on them.

- Justice Scalia's rejoinder: "We don't have any urns here."

- Should model this combinatorially
  - Ball draws not independent trials (balls not replaced)
- Exact solution:

  P(draw 12 black balls) = $\binom{940}{12} \Big/ \binom{1000}{12} \approx 0.4739$

  P(draw ≥ 1 red ball) = 1 – P(draw 12 black balls) ≈ 0.5261

- Approximation using Binomial distribution
  - Assume P(red ball) constant for every draw = 60/1000
  - X = # red balls drawn. X ~ Bin(12, 60/1000 = 0.06)
  - P(X ≥ 1) = 1 – P(X = 0) ≈ 1 – 0.4759 = 0.5240

In Breyer's description, should actually expect just <u>over half</u> of juries to have at least one black person on them

X a random variable, A an event.

The expectation of X conditioned on A is defined as:

$$E(X|A) = \sum_{a \in \mathcal{A}} a \cdot Pr(X = a|A)$$

Roll a fair die. Let random variable R be the number showing

$$E(R|R \geq 4) = \sum_{i=1}^{6} i \cdot Pr(R = i|R \geq 4)$$

X random variable on probability space Ω

$A_1, A_2, \ldots, A_k$ partition of Ω

$$E(X) = \sum_{i=1}^{k} E(X|A_i) \cdot Pr(A_i)$$

$$E(X) = \sum_{a \in \mathcal{A}} a \cdot Pr(X = a)$$

$$= \sum_{a \in \mathcal{A}} a \sum_{i=1}^{k} Pr(X = a|A_i) Pr(A_i)$$

$$= \sum_{i=1}^{k} Pr(A_i) \sum_{a \in \mathcal{A}} a \cdot Pr(X = a|A_i)$$

LTE: $$E(X) = \sum_{i=1}^{k} E(X|A_i) \cdot Pr(A_i)$$

computer crashes with probability p each hour

X: # hours till it crashes.

E(X) = p E(X | crashes during first hour) +
(1-p) E(X | doesn't crash during first hour)

LTE:

$$E(X) = \sum_{i=1}^{k} E(X|A_i) \cdot Pr(A_i)$$

computer crashes with probability p each hour

X: # hours till it crashes.

E(X) = p E(X | crashes during first hour) +
(1-p) E(X | doesn't crash during first hour)
= p + (1-p) [1 + E(X)]

Solve for E(X) E(X) = 1/p

A prisoner is trapped in a cell containing 3 doors. The first door leads to a tunnel that returns him to his cell after an amount of time that is Poisson(5). The second leads to a tunnel that returns him to his cell after an amount of time that Geometric(1/3). The third door leads to freedom after 1 day of travel.

If it is assumed that the prisoner will always select doors 1, 2 and 3 with respective probabilities 0.5, 0.3 and 0.2, what is the expected number of days until the prisoner reaches freedom?

Often care about 2 (or more) random variables *simultaneously*

measured X = height and Y = weight

X = cholesterol and Y = blood pressure

X_1, X_2, X_3 = work loads on servers A, B, C

Joint probability mass function:

$f_{XY}(x, y) = P(X = x \ \& \ Y = y)$

Joint cumulative distribution function:

$F_{XY}(x, y) = P(X \leq x \ \& \ Y \leq y)$

Two joint PMFs

W \ Z	1	2	3
1	2/24	2/24	2/24
2	2/24	2/24	2/24
3	2/24	2/24	2/24
4	2/24	2/24	2/24

X \ Y	1	2	3
1	4/24	1/24	1/24
2	0	3/24	3/24
3	0	4/24	2/24
4	4/24	0	2/24

P(W = Z) = 3 * 2/24 = 6/24

P(X = Y) = (4 + 3 + 2)/24 = 9/24

Can look at arbitrary relationships between variables this way

Two joint PMFs

W \ Z	1	2	3	$f_W(w)$
1	2/24	2/24	2/24	6/24
2	2/24	2/24	2/24	6/24
3	2/24	2/24	2/24	6/24
4	2/24	2/24	2/24	6/24
$f_Z(z)$	8/24	8/24	8/24	

X \ Y	1	2	3	$f_X(x)$
1	4/24	1/24	1/24	6/24
2	0	3/24	3/24	6/24
3	0	4/24	2/24	6/24
4	4/24	0	2/24	6/24
$f_Y(y)$	8/24	8/24	8/24	

Marginal distribution of one r.v.: sum over the other:

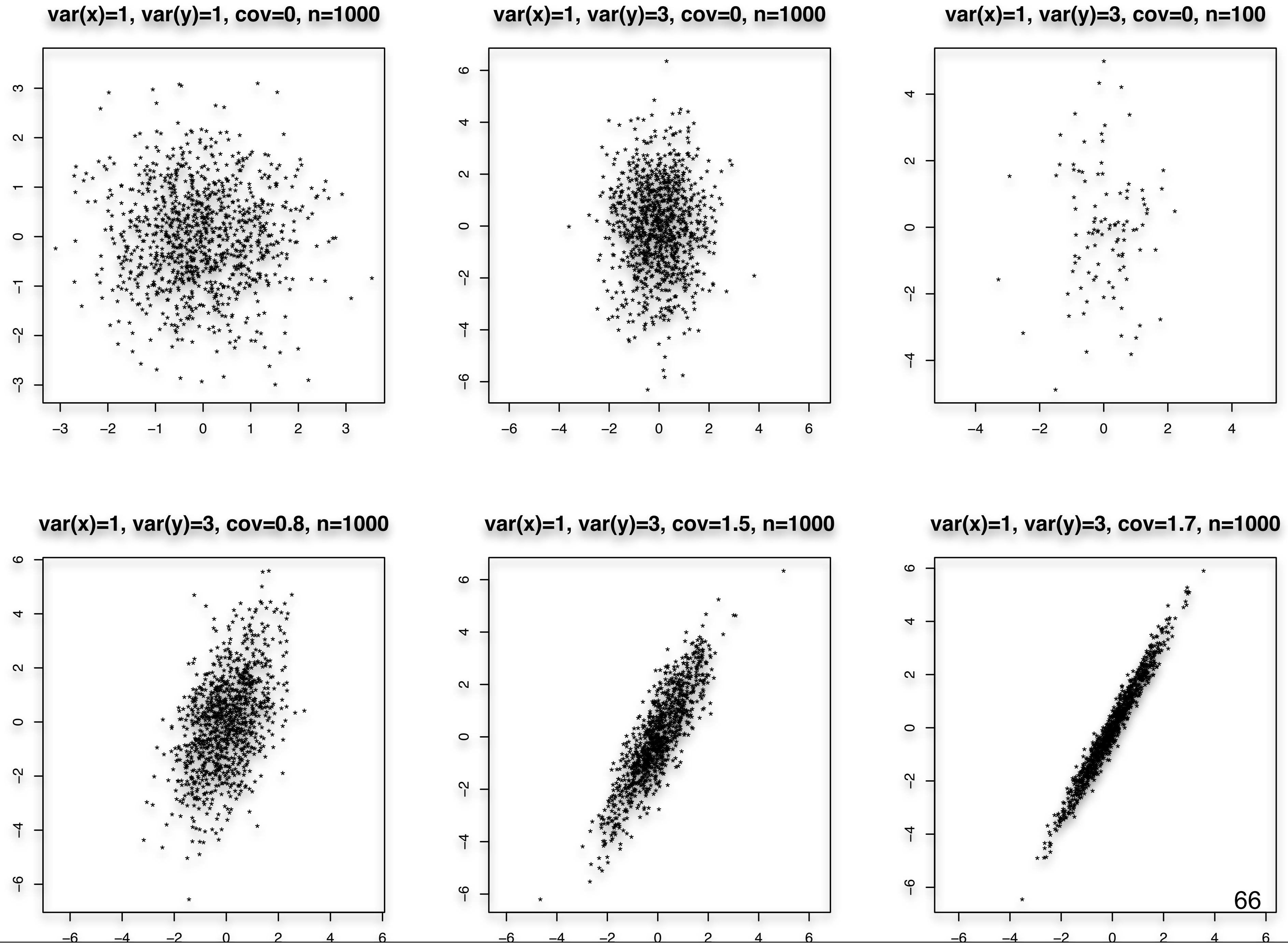
$f_Y(y) = \Sigma_x f_{XY}(x,y)$ $\quad$ $f_X(x) = \Sigma_y f_{XY}(x,y)$

Question: Are W & Z independent? Are X & Y independent?

sampling from a (continuous) joint distribution

Top row; independent variables

bottom row: dependent variables



X is equally likely to be (1, 2, 3)

Y is always = X+ 3

What is the joint distribution?

A function g(X,Y) defines a new random variable.

Its expectation is:

$$E[g(X, Y)] = \Sigma_x \Sigma_y\, g(x, y) f_{XY}(x,y)$$

Expectation is linear. I.e., if g is linear:

$$E[g(X, Y)] = E[a\,X + b\,Y + c] = a\,E[X] + b\,E[Y] + c$$

Example:

g(X,Y) = 2X-Y

E[g(X,Y)] = 72/24 = 3

E[g(X,Y)] = 2•2.5 - 2 = 3

X \ Y	1	2	3
1	1 • 4/24	0 • 1/24	-1 • 1/24
2	3 • 0/24	2 • 3/24	1 • 3/24
3	5 • 0/24	4 • 4/24	3 • 2/24
4	7 • 4/24	6 • 0/24	5 • 2/24

RV: a numeric function of the outcome of an experiment

Probability Mass Function p(x): prob that RV = x; $\Sigma p(x)=1$

Cumulative Distribution Function F(x): *probability that RV ≤ x*

Expectation:

of a random variable: $E[X] = \Sigma_x \, xp(x)$

of a function: if $Y = g(X)$, then $E[Y] = \Sigma_x \, g(x)p(x)$

linearity:

$E[aX + b] = aE[X] + b$

$E[X+Y] = E[X] + E[Y]$; even if *dependent*

this interchange of "order of operations" is quite special to linear combinations. E.g. $E[XY] \neq E[X]*E[Y]$, in general (but see below)

Variance:

$Var[X] = E[(X-E[X])^2] = E[X^2] - (E[X])^2]$

Standard deviation: $\sigma = \sqrt{Var[X]}$

$Var[aX+b] = a^2 Var[X]$

If X & Y are *independent*, then

$E[X \cdot Y] = E[X] \cdot E[Y]$;

$Var[X+Y] = Var[X]+Var[Y]$

(These two equalities hold for *indp* rv's; but not in general.)

Important Examples:

Bernoulli: $P(X=1) = p$ and $P(X=0) = 1-p$ $\mu = p, \quad \sigma^2 = p(1-p)$

Binomial: $P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}$ $\mu = np, \sigma^2 = np(1-p)$

Poisson: $P(X = i) = e^{-\lambda} \frac{\lambda^i}{i!}$ $\mu = \lambda, \quad \sigma^2 = \lambda$

Bin(n,p) $\approx$ Poi(λ) where $\lambda = np$ fixed, $n \to \infty$ (and so $p=\lambda/n \to 0$)

Geometric $P(X=k) = (1-p)^{k-1}p$ $\mu = 1/p, \sigma^2 = (1-p)/p^2$

Many others, e.g., hypergeometric