Chapter 4 Expectations of functions of RVs

Example: Toss a die and win X^2 if X is # of spots.

$$E(X^2) = \sum_x x^2 p_X(x) = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + \dots + 6^2 \cdot \frac{1}{6}$$

$$= 15\tfrac{1}{6}$$

Note $E(X^2) \neq (E(X))^2$

$$(E(X))^2 = \left(1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \dots + 6 \cdot \frac{1}{6}\right)^2$$

$$= (3.5)^2 = 12\tfrac{1}{4}$$

In general: $E(g(X)) = \sum_x g(x)\, p_X(x)$

or

$= \int g(x)\, f_X(x)\, dx$

Works for function of random vectors:

Let $Y = g(X_1, X_2, \ldots, X_k)$

with pmf $p(x_1, x_2, \ldots, x_k)$

then $E(Y) = E(g(X_1, \ldots, X_k)) = \sum_{x's} g(x_1, \ldots, x_k)\, p(x_1, \ldots, x_k)$

or $$E(Y) = \int \left[\int \cdots \int g(x_1, \ldots, x_k) f(x_1, \ldots, x_k)\, dx_k \cdots dx_2 \right] dx_1$$

if the integral (sum) with $|g|$ converges.

---

p.124 Corollary A

If X & Y are independent, then

$$E(g(X)h(Y)) = E(g(X)) \times E(h(Y))$$

provided ↑ and ↑ exist.

Special case: If X & Y are independent

$$E(XY) = E(X)E(Y)$$

if ↑ and ↑ exist.

Beware: 1) not true in general

2) converse not true

Linear combinations of R.V.'s

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \cdots + b_nX_n$$

then $E(Y) = a + b_1E(X_1) + b_2E(X_2) + \cdots + b_nE(X_n)$

if ↑ $\cdots$ exist.

Example A (p. 126)

What is $E(Y)$ if $Y \sim Bin(n, p)$?

$$E(Y) = \sum_y \binom{n}{y} y\, p^y (1-p)^{n-y}$$

Easier: Use fact that $Y = X_1 + \cdots + X_n$ where X_i's are indep. $Bin(1, p)$ and $E(X_i) = p$

So $E(Y) = \sum_{i=1}^{n} E(X_i) = np$

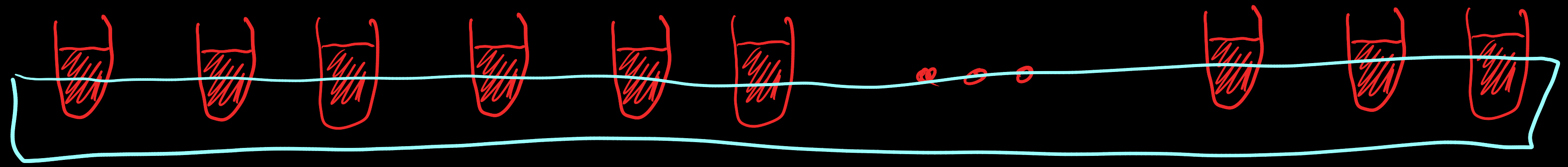
Example C p. 128

Group testing

- n blood samples tested for rare disease

- If you test each individually will need n tests.

Alternative:



take 1/2 of each sample and combine:

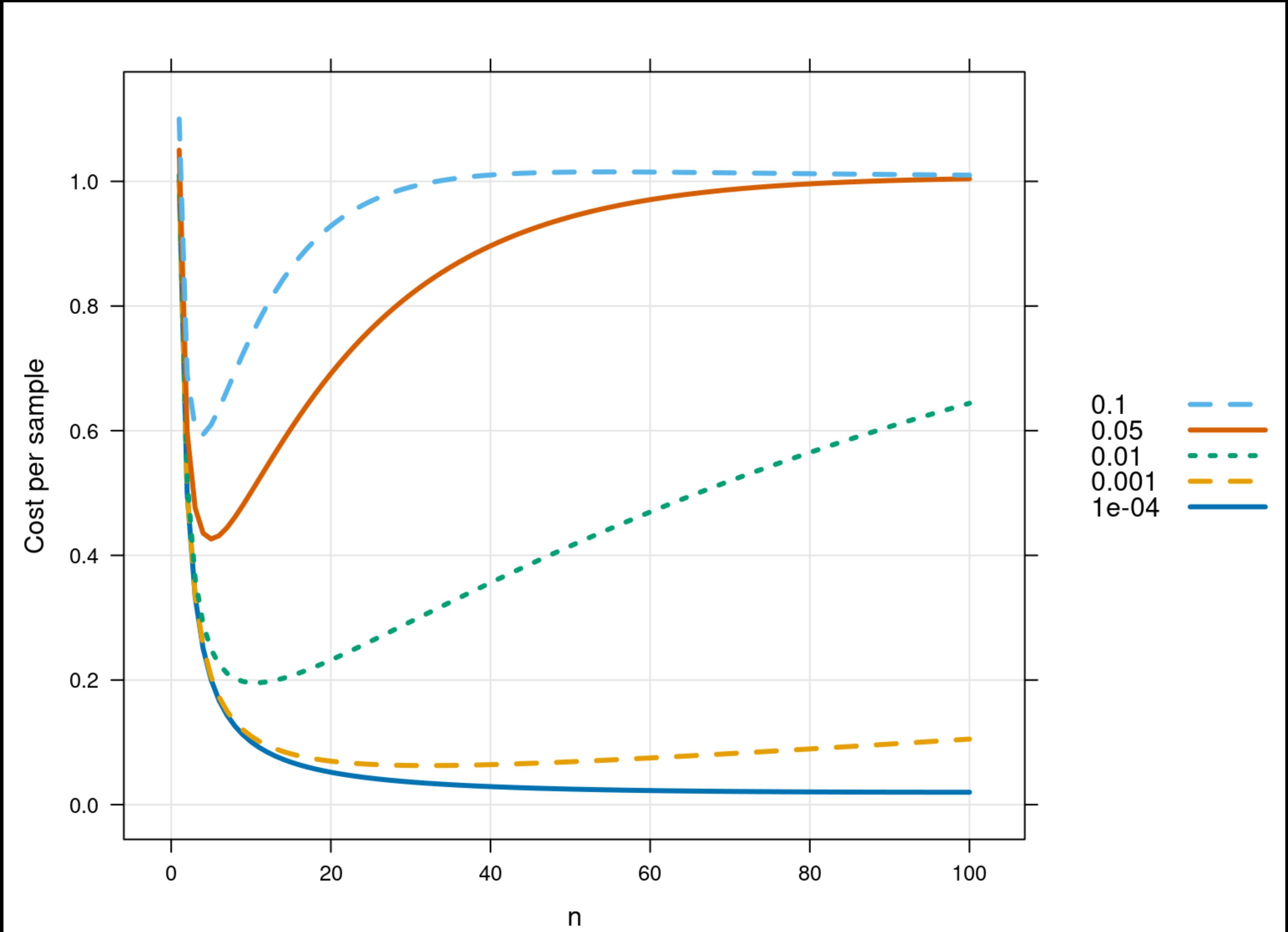
Then test.

If negative - done - all okay

If positive - test n samples.

Let p = probability of a positive.

$$E(\text{Tests}) = 1 \times \underbrace{(1-p)^n}_{\text{pr all negative}} + (n+1)\underbrace{\left(1-(1-p)^n\right)}_{\text{prob. at least 1 positive}}$$



1.0
0.8
0.6
0.4
0.2
0.0
Cost per sample
0
20
40
60
80
100
n
0.1
0.05
0.01
0.001
1e-04

```r
df <- expand.grid(n = 1:100, p = c(.0001,.001, .01, .05, .1))
head(df)
dim(df)
df <- within(df,
             {
               Etests <- 1* (1-p)^n + (n + 1) * (1 - (1 - p)^n)
               Cost_per_sample <- Etests/n
             }
             )

library(latticeExtra)
trellis.par.set(superpose.line = list(lwd=3, lty = 1:3))
xyplot(Cost_per_sample ~ n, df,
       groups = p, type = 'l',
       ylab = "Cost per sample",
       auto.key = list(reverse.rows = T)) +
  layer_(panel.grid(h=-1, v = -1))
```

Example D: Illustrates how

$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$$

does not require independence

DNA Sequences: formed from 4 letters A C T G

If random & each letter with = prob.

$$\underbrace{\text{ATCAATCGAGT} \cdots \text{TAA}}$$

- Suppose length = N
- and each letter has $p = 1/4$

How many ATGC do you expect?

Let I_n = event ATGC starts at position n

$$P(I_n) = \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} = \frac{1}{256}$$

$$= E(I_n)$$

if $I_n = \begin{cases} 1 & \text{if ATGC starts at position } n \\ 0 & \text{otherwise} \end{cases}$

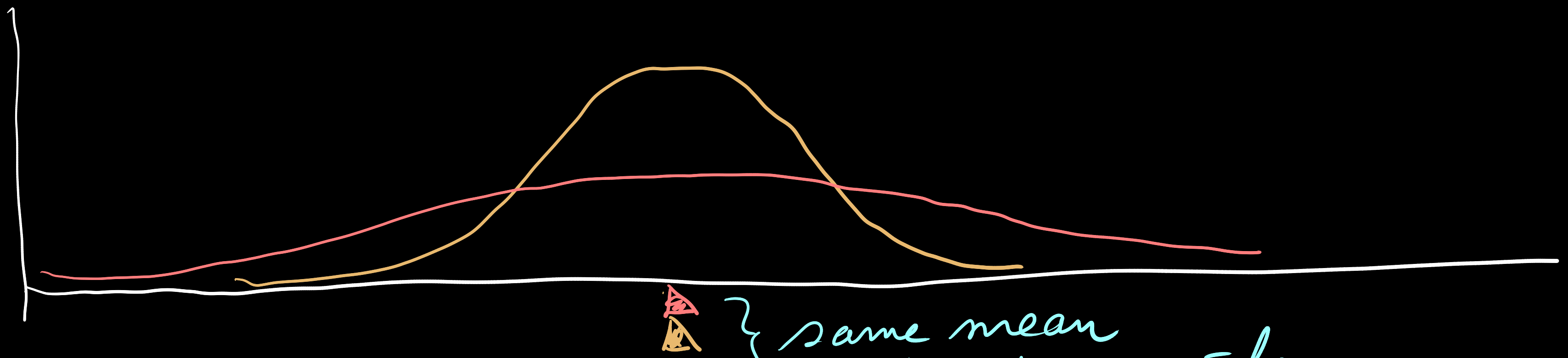
$$E(\#\text{ of sequences}) = E\left(\sum_{n=1}^{N-3} I_n\right) = \sum_{n=1}^{N-3} E(I_n)$$

$$= (N-3) \times \frac{1}{256}$$

Variance and Standard Deviation

Mean = "location parameter" ← one of many e.g. median, min, max

Next we need a "spread" parameter



} same mean
different spreads

Variance: Average squared distance from the mean

$$\mathrm{Var}(X) = E\left[(X-\mu_X)^2\right]$$ if E exists. in squared units

$$SD(X) = \sqrt{\mathrm{Var}(X)}$$ in original units

Facts about variance.

$$\mathrm{Var}(X) = E(X^2) - (E(X))^2$$

also a useful way to calculate Var(X)

Proof

$$\begin{aligned} E\left[(X-\mu_X)^2\right] &= E\left(X^2 - 2\mu_X X + \mu_X^2\right) \\ &= E(X^2) - 2\mu_X E(X) + \mu_X^2 \\ &= E(X^2) - 2E(X)E(X) + (E(X))^2 \\ &= E(X^2) - (E(X))^2 \end{aligned}$$

Corollary: $E(X^2) = E(X)^2$ iff $Var(X) = 0$

Fact: $Var(X) = 0$ iff X is a constant

i.e. $P(X = c) = 1$

Fact: If $Var(X) < \infty$ (i.e. exists)

and $Y = a + bX$

then $Var(Y) = b^2 \, Var(X)$

Chebyshev's Inequality : Prob & spread

Let X have mean μ and variance σ^2

Then for any $t \geq 0$:

$$P(|X-\mu| \geq t) \leq \sigma^2/t^2$$

Proof: Use Markov's inequality on $Y = (X-\mu)^2$