transmit image, ... color white or gray
{0 else

Assume X_i are iid (correlated more reasonable)

with $P(X_i=1) = P$(pixel color is correct for received image) $= 0.99$

Let W = # correct pixels so $W \sim bin(64, 0.99)$,

P(2 or more received pixels are incorrect) $= 1 - P$(0 or 1 pixels are incorrect)

Y = # incorrect $\sim bin(64, 0.01)$, $P_k = \binom{n}{k} p^k (1-p)^{n-k}$

$= 1 - P(Y=0) - P(Y=1) = 1 - \binom{64}{0}(.01)^0(.99)^{64} - \binom{64}{1}(.01)(.99)^{63}$

$= 1 - .5256 - .3398 = 0.1346$

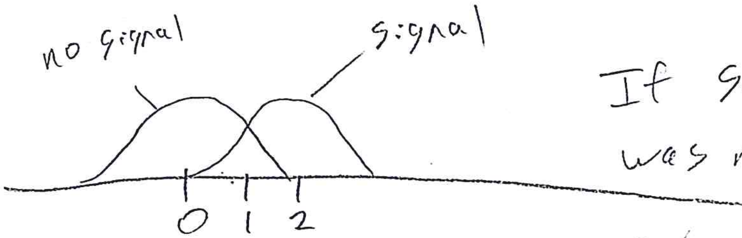
$= 1 - P(W=64) - P(W=63) = 1 - \binom{64}{64}(.99)^{64}(.01)^0 - \binom{64}{63}(.99)^{63} .01$

Image processing is much more complex with "error correcting" and Bayesian methods.

not important ↑

ex] Signal processing (Ross covers -1 $+1$ signals (telegraph) several times.)

If a signal is not sent the receiver get RV $X_i \sim N(0,1)$. If the signal is sent $X_i \sim N(2,1)$. Decide signal sent if $X_i \geq 1$
not sent if $X_i < 1$



If signal is sent, P(decide signal was not sent) $= P(X_i < 1) =$

$P\left(Z < \frac{1-2}{1}\right) = P(Z < -1) = 0.1587$

If signal was not sent, P(decide signal was sent) $= P(Z > 1) = 0.1587$.

1) Simulation uses random number generation to produce pseudo random variables (or random vectors) from a specified distribution.

p645-6

2) Suppose you want to estimate $E[g(\underline{X})]$ and you can generate $\underline{X}_1, \ldots, \underline{X}_n \sim$ iid from the same dist as $\underline{X}_1$.

Let $W_i = g(\underline{X}_i)$ for $i = 1, \ldots, n$.

Then $\overline{W} = \frac{1}{n}\sum_{i=1}^{n} W_i = \frac{1}{n}\sum_{i=1}^{n} g(\underline{X}_i) \to E[g(\underline{X})]$ by SLLN.

This technique is called Monte Carlo simulation.

3) If you can generate $U_1, \ldots, U_n \overset{iid}{\sim} U[0,1]$, you can generate $X_1, \ldots, X_n \overset{iid}{\sim} X$ for many dist's.

4) Idea: generate pseudo random (numbers) variables $X_1, \ldots, X_n$ that act like a sample $X_1, \ldots, X_n$ iid from a brand name distribution.

ex) R: z ← runif(100) makes 100 pseudo iid U(0,1) RVs

$\frac{1}{n}\sum X_i = \overline{X} =$ mean(z) = 0.5263 ($E(X) = 0.5$)

$\frac{1}{n-1}\sum(X_i - \overline{X})^2 = S^2 = \hat{\sigma}^2 =$ var(z) = 0.08249 ≈ ($V(X) = \frac{1}{12} \approx 0.08333$)

5.5 p 649–650 Inversion method: 480 67

Suppose X has a pdf $f(x)$ and increasing cdf $F(x)$.

Let $F^{-1}(u)$ be the inverse of F: solve

$u = F(x)$ for $x = F^{-1}(u)$. ($F^{-1}(u)$ will often

be given and see exam 3 review 93) for

$F^{-1}(u)$ for some brand name dist's.)

Suppose $U_1, \ldots, U_n$ are n pseudo $U(0,1)$ RVs.

Then $X_1, \ldots, X_n$ are pseudo RVs from the dist of X

if $X_i = F^{-1}(U_i)$.

See HW 11 #2

ex] $Z \in$ runif(5)

u	0.569	.883	.174	.887	.572
$F^{-1}(u) = x$	0.842	2.145	0.191	2.180	0.849

Let $X \sim Exp(1)$ so $F(x) = 1 - e^{-x} = u$. Then

$e^{-x} = 1-u$ and $-x = \log(1-u)$.

Hence $x = -\log(1-u) = F^{-1}(u)$.

Let $X = F^{-1}(U)$.

i] Suppose $x_i = F^{-1}(u_i)$. Then $u_i = F(x_i)$.

$P(X \le x_i) = P(F^{-1}(U) \le x_i) = P(U \le F(x_i))$

$= P(U \le u_i) = u_i$. So x_i is the $100 u_i$ th percentile of x! $P(X \le x_i) = v_i$.

(67.5)

7] The 100αth pop percentile of X is x_α

where $P(X \le x_\alpha) = \alpha$ and $0 < \alpha < 1$.

Percentiles are also called quantiles.

p54
8] Arrange data $x_1, \ldots, x_n$ in ascending order to get the order statistics $x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)}$.

$\uparrow$ min $\qquad$ $\uparrow$ max

Then $x_{(i)}$ is the ith order statistic.

9] Let $\lfloor x \rfloor$ = greatest integer $\le x$, ie $\lfloor 7.7 \rfloor = 7$.

For hand calculations, let $j = \lfloor (n+1)\alpha \rfloor$.

The estimator of x_α is the sample percentile

$\hat{x}_\alpha = x_{(j+1)}$ (take $\hat{x}_\alpha = x_{(n)}$ if $j = n$).

10] R uses $(1-h)\, x_{(j)} + h\, x_{(j+1)} = \hat{x}_\alpha$ for some h

where $0 \le h \le 1$.

see HW11 4b)

ex] $1, 2, 3, 4, 5, 6, 7, 8, 9, 10$ $\leftarrow x_{(1)}, \ldots, x_{(n)}$

$x_{(i)} = i$

Find $\hat{x}_{.8}$. Soln $j = \lfloor (n+1)\alpha \rfloor = \lfloor 11(.8) \rfloor = \lfloor 8.8 \rfloor = 8$

So $\hat{x}_{.8} = x_{(j+1)} = x_{(9)} = 9$

11) A population 100 $(1-\alpha)$% prediction interval (PI) $[L, U]$ satisfies $F(U) - F(L) \geq 1-\alpha$. A large sample 100 $(1-\alpha)$% PI $[\hat{L}_n, \hat{U}_n]$ satisfies

$P[\hat{L}_n \leq X \leq \hat{U}_n] \to 1-\delta \geq 1-\alpha$ as $n \to \infty$. The population shorth $[L_S, U_S]$ is the shortest population 100 $(1-\alpha)$% PI.

12) Consider intervals that contain c cases

$[X_{(1)}, X_{(c)}]$, $[X_{(2)}, X_{(c+1)}]$, ..., $[X_{(n-c+1)}, X_{(n)}]$.

Compute $X_{(c)} - X_{(1)}$, $X_{(c+1)} - X_{(2)}$, ..., $X_{(n)} - X_{(n-c+1)}$.

Then shorth(c) = $[X_{(d)}, X_{(d+c-1)}]$ is the closed interval with shortest length. If $\frac{c}{n} \to 1-\alpha$, the shorth (c) PI estimates the population shorth.

Ex] Let $c = 4$ Data below has $n = 7$. See HW 4a)

0, 1, 3, 6, 9, 10, 11

$6 = 6 - 0$

$8 = 9 - 1$

$7 = 10 - 3$

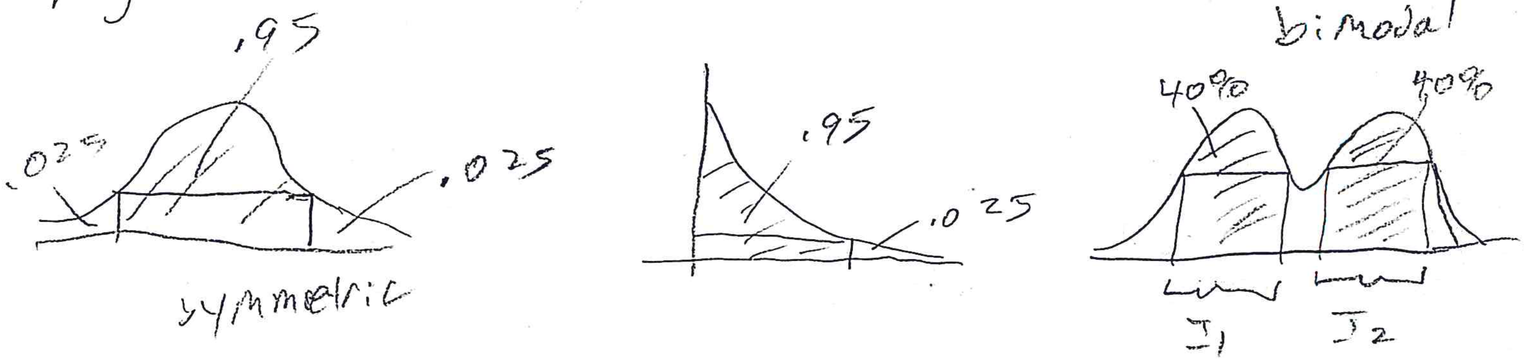
$5 = 11 - 6$ ← 5 is shortest length

intervals containing 4 cases

shorth(4) = $[6, 11]$.

13) The highest $100(1-\alpha)\%$ density region of a pdf is found by moving a horiz line down from the top of the pdf so that the line intersects the pdf (or border) at one or more intervals and the sum of the areas under the pdf corresponding to the intervals $= 1-\alpha$. The pdf can't have a positive flat interval (eg $U(a,b)$).

For a unimodal pdf, the pop shorth = highest density region.



14) Let $x_1, \ldots, x_n$ be pseudo RVs from RV X.

checks: $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \approx E(X)$

$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \approx V(X)$ $\quad$ n large

$\hat{x}_{(.975)} - \hat{x}_{.025} \approx x_{.975} - x_{.025}$

shorth$[n(0.95)] \approx$ pop shorth (95%)

ex) z <- runif(100), x <- -log(1-z) $\approx EXP(1)$

mean(x) $= \bar{x} = .9755$ $\quad$ $E(X) = 1$
var(x) $= s^2 = .9092$ $\quad$ $V(X) = 1$