Lecture 18: Feb 18, Indicator random variables: Ross 7.2,7.3

18.1: Counting events using indicator random variables: Ross P.340

Let A_1, ..., A_n be events (any subsets of Ω). Let $X_i = 1$ if A_i occurs, and $X_i = 0$ otherwise.

Note $\mathrm{E}(X_i) \ = \ P(A_i)$, and $\prod_{i=1}^k X_i \ = \ 1$ if and only if $X_1 = X_2 = ... = X_k = 1$, i.e. $\cap_1^k A_i$ occurs.

Also $1 \ - \ \prod_{i=1}^k (1 - X_i) \ = \ 1$ if and only if some $X_i = 1, \ i = 1,..,k$ i.e. $\cup_1^k A_i$ occurs.

$$\begin{aligned}
P(A \cup B \cup C) &= \mathrm{E}(1 - (1 - X_A)(1 - X_B)(1 - X_C)) \\
&= \mathrm{E}(X_A + X_B + X_C - X_A X_B - X_A X_C - X_B X_C + X_A X_B X_C) \\
&= P(A) + P(B) + P(C) \ - P(A \cap B) \ - P(A \cap C) \ - P(B \cap C) + P(A \cap B \cap C)
\end{aligned}$$

18.2: Expectation of counts of events: Ross P.347-348

Let A_i and X_i be as above, and let $Y \ \equiv \ \sum_{i=1}^n X_i$ be the number of events that occur. Let $p_i = P(A_i) = \mathrm{E}(X_i)$.

Note $\mathrm{var}(X_i) = p_i(1 - p_i)$.

$$\begin{aligned}
\mathrm{E}(Y) &= \mathrm{E}(\sum_{i=1}^n X_i) \ = \ \sum_{i=1}^n \mathrm{E}(X_i) \ = \ \sum_{i=1}^n p_i \\
\mathrm{var}(Y) &= \mathrm{var}(\sum_{i=1}^n X_i) \ = \ \sum_{i=1}^n \mathrm{var}(X_i) \ + \ 2 \sum \sum_{i<j} \mathrm{cov}(X_i, X_j) \\
&= \sum_{i=1}^n p_i(1 - p_i) \ + \ 2 \sum \sum_{i<j} (P(A_i \cap A_j) - p_i p_j) \\
&= \sum_{i=1}^n p_i \ - \ (\sum_i p_i)^2 \ + \ 2 \sum \sum_{i<j} P(A_i \cap A_j) \ = \ \mathrm{E}(Y) \ - \ (\mathrm{E}(Y))^2 + 2 \sum \sum_{i<j} P(A_i \cap A_j)
\end{aligned}$$

18.3: Mean and variance of a hypergeometric random variables: Ross P.350

We have N fish in a pond. m are red and $N - m$ are blue. We sample n fish.

Let Y be the number of red fish in the sample. Let A_i be event that ith fish sampled is red.

Consider the ith fish sampled; every fish has the same probability to be the $i$the fish; $\mathrm{E}(X_i) = P(A_i) = m/N$.

So $\mathrm{E}(Y) \ = \ \sum_{i=1}^n \mathrm{E}(X_i) \ = \ nm/N$. Now consider jth fish: $P(X_j = 1 \mid X_i = 1) \ = \ (m - 1)/(N - 1)$.

$\mathrm{E}(X_i X_j) \ = \ m(m - 1)/N(N - 1)$. So, from 18.2:

$$\begin{aligned}
\mathrm{var}(Y) &= nm/N \ - \ (nm/N)^2 \ + \ 2\binom{n}{2} m(m-1)/N(N-1) \ = \ \frac{nm}{N}\left(1 - \frac{nm}{N} + \frac{(n-1)(m-1)}{N-1}\right) \\
&= n(m/N)(1 - (m/N) \ - \ nm(N - m)/N^2(N - 1)
\end{aligned}$$

Note this is smaller than the Binomial variance with the same $\mathrm{E}(X_i) = p = m/N$.

18.4: Variances and covariances in a Multinomial distribution: Ross P.364.

Let $(Y_1, ..., Y_K)$ be counts of outcomes type $k = 1, 2, ...K$, in a multinomial sample size n.

Let $X_{ki} = 1$ if ith outcome is type k, 0 otherwise. $\mathrm{E}(X_{ki}) = p_k$.

$Y_k \ = \ \sum_{i=1}^n X_{ki}$, and outcomes i and j are independent.

So Y_k is Binomial (n, p_k); $\mathrm{E}(Y_k) = np_k$, $\mathrm{var}(Y_k) = np_k(1 - p_k)$.

However, for outcome i, cannot be both k and k' so $X_{ki} X_{k'i} \equiv 0$. So we have

$$\begin{aligned}
\mathrm{cov}(Y_k, Y_{k'}) &= \mathrm{cov}(\sum_{i-1}^n X_{ki}, \sum_{j=1}^n X_{k'j}) \ = \ \sum_i \sum_j \mathrm{cov}(X_{ki}, X_{k'j}) \\
&= \sum_i \mathrm{cov}(X_{ki}, X_{k'i}) \ = \ \sum_i (0 \ - \ p_k p_{k'}) \ = \ -np_k p_{k'}.
\end{aligned}$$

Lecture 19: Feb 20, Examples of expectations, covariances and indicator r.v.

1. Multinomial example

12 students go to give blood. They are from a population in which the frequency of the 4 types A, B, AB and O are 0.3, 0.2, 0.1 and 0.4, respectively.

(i) What is the expected number of A blood-type students? What is the variance?

(ii) What is the covariance between the number of A-type and AB type students?

(iii) What is the correlation between the number of A-type and not-A-type students?

(iv) What is the variance of the number of students who are not A-type?

Important note: In the multinomial, the blood types of different students are independent. Whether student i is A is independent of whether student j is AB. The negative cobvariances in the total count of A and AB students comes about because the *same* student cannot be both A and AB. (cf. question 3).

2. Ch 7 # 22, and ST #4.

(a) We want to find the expected number of tosses to get all faces of die.

Suppose I have seen $6 - k$ different faces already, so I have k left to get.

(i) What is the probability I get a new face on the next toss?

(ii) What is the expected number of tosses to the next new face?

(iii) What is the expected total to get all faces? (Answer: 14.7)

(b) We want to find the expected number of 6's in these tosses.

(i) Let N_k be expected number of k's; note $N_1 = N_2 = ... = N_6$.

(ii) So what is the answer to (b)?

3. Dependence in sampling; Let $X_i = 1$ if ith sampled object specified type k (e.g. red), 0 otherwise.

(a) In a multinomial/binomial model (sampling with replacement), the probability of type k is p_k on every trial. X_i and X_j are independent: $\mathrm{E}(X_i X_j) = p_k^2$. $\mathrm{cov}(X_i, X_j) = 0$.

(b) In a hypergeometric model (sampling without replacement), the probability of a red fish is decreased if we have already sampled red fish. Show $\mathrm{cov}(X_i, X_j) = -m(N-m)/N(N-1)$.

(c) In a Polya urn model, we can increase the probability of a type k outcome type k with each type k sampled. We have m red balls, and $N - m$ blue balls in an urn, and an unlimited extra pile of balls. Each time we select a red ball we replace it, *and add an additional red ball.* Likewise, if we sample a blue ball, we add an *additional* blue ball.

Let $X_k = 1$ if kth ball drawn is red, and 0 otherwise.

(i) What is $\mathrm{E}(X_1)$?

(ii) What is $\mathrm{E}(X_2)$?

(iii) What do you conclude about $\mathrm{E}(X_i)$ for any i?

(iii) What is $\mathrm{E}(X_1 X_2)$, and $\mathrm{cov}(X_1, X_2)$.

(iv) What is the probability of the sequence $BBRRRBB$?

(v) What is the probability of the sequence $RBBRBRB$?

(vi) What do you conclude about $\mathrm{cov}(X_i, X_j)$ for any i,j?

(vii) How would you find the mean and variance of the number of red balls in n draws?

Lecture 20: Feb 23: More mgf's; the bivariate Normal dsn

20.1 The mgf of a Normal random variable: left from Lecture 17

Normal: $N(0,1)$: note $-\frac{1}{2}x^2 + tx \;=\; -\frac{1}{2}(x-t)^2 + \frac{1}{2}t^2$

$$\mathrm{E}(e^{tX}) \;=\; (1/\sqrt{2\pi})\int_{-\infty}^{\infty} \exp(-\frac{1}{2}x^2 + tx)\;dx \;=\; (\exp(\frac{1}{2}t^2)/\sqrt{2\pi})\int_{-\infty}^{\infty} \exp(-\frac{1}{2}(x-t)^2)\;dx \;=\; \exp(\frac{1}{2}t^2)$$

Let $Y = aX + b$: $M_Y(t) = \mathrm{E}(\exp((aX+b)t)) \;=\; e^{bt}\mathrm{E}(\exp((at)X)) \;=\; e^{bt}M_X(at)$.

Let $X \sim N(\mu, \sigma^2)$, so $X = \sigma Z + \mu$ where $Z \sim N(0,1)$, so $M_X(t) = \exp(\mu t \;+\; \frac{1}{2}\sigma^2 t^2)$.

Then $Y = aX + b$ gives

$M_Y(t) \;=\; e^{bt}\exp(\mu(at) \;+\; \frac{1}{2}\sigma^2(at)^2) \;=\; \exp((a\mu + b)t \;+\; \frac{1}{2}(a^2\sigma^2)t^2)$ so $Y \sim N(a\mu + b, a^2\sigma^2)$.

Sum of independent Normals (any mean/variance) is Normal; $X \sim N(\mu, \sigma^2)$, $Y \sim N(\nu, \tau^2)$:

$M_{X+Y}(t) \;=\; M_X(y)M_Y(t) \;=\; \exp(\mu t \;+\; \frac{1}{2}\sigma^2 t^2)\exp(\nu t \;+\; \frac{1}{2}\tau^2 t^2) \;=\; \exp((\mu+\nu)t + \frac{1}{2}(\sigma^2 + \tau^2))$,

so $X + Y \sim N(\mu + \nu, \sigma^2 + \tau^2)$.

Hence any linear combination of independent Normals is also Normal.

20.2 The bivariate Normal via independent standard Normals

Let Z and W be independent $N(0,1)$, $X \;=\; \mu + aZ + bW$, $Y \;=\; \nu + cZ + dW$.

Then X and Y are said to be jointly (bivariate) Normal.

Note $X \sim N(\mu, \sigma^2)$ with $\sigma^2 = a^2 + b^2$. Note $Y \sim N(\nu, \tau^2)$ with $\tau^2 = c^2 + d^2$.

Also $\mathrm{cov}(X,Y) \;=\; \mathrm{E}((aZ + bW)(cZ + dW)) \;=\; (ac + bd), or \rho(X,Y) = (ac+bd)/\sigma\tau$.

The pdf of (X, Y) is given by Ross in Ch 6. Pp.294-5 – a nasty mess.

20.3 The mgf of bivariate Normal (X, Y)

Define $M_{X,Y}(s,t) \;=\; \mathrm{E}(\exp(sX + tY))$. Note $M_X(s) = M_{X,Y}(s, 0)$, $M_Y(t) = M_{X,Y}(0,t)$.

$$\begin{aligned} M_{X,Y}(s,t) &= \mathrm{E}(\exp(s\mu + asZ + bsW + t\nu + ctZ + dtW)) \;=\; \exp(s\mu + t\nu)M_Z(as+ct)M_W(bs+dt) \\ &= \exp(s\mu + t\nu + \frac{1}{2}(as+ct)^2 + \frac{1}{2}(bs+dt)^2) \;=\; \exp(s\mu + +t\nu + \frac{1}{2}\sigma^2 s^2 + \frac{1}{2}\tau^2 t^2 \;+\; (ac+bd)st) \\ &= M_X(s)\; M_Y(t)\; \exp(\rho\sigma\tau st) \end{aligned}$$

20.4 Independence if and only if $\rho = 0$

Recall in general, $\rho(X,Y) = 0$ does **not** imply X, Y independent.

Recall that the mgf **uniquely** identifies the pdf; the same is true for our joint mgf $M_{X,Y}(s,t)$ and the joint pdf $f_{X,Y}(x,y)$.

Now suppose $\rho = 0$ in our bivariate Normal; i.e. $ac + bd = 0$.

Then from 20.3: $M_{X,Y}(s,t) \;=\; M_X(s)M_Y(t)$, but this is the joint mgf of two independent Normals $X \sim N(\mu, \sigma^2)$ and $Y \sim N(\nu, \tau^2)$, so by uniqueness, X and Y are two **independent** Normals.

That is **for Normal random variables**, $\rho = 0$ **does** imply independence.

The same can be seen from the very messy joint density given by Ross on P.294:

$$\begin{aligned} f_{X,Y}(x,y) \;=\; (2\pi\sigma\tau\sqrt{1-\rho^2})^{-1}\exp\Bigg(-\frac{1}{2(1-\rho^2)}\bigg((\frac{x-\mu}{\sigma})^2 \\ -2\rho\frac{(x-\mu)(y-\nu)}{\sigma\tau} \;+\; (\frac{y-\nu}{\tau})^2\bigg)\Bigg) \end{aligned}$$

on $-\infty < x < \infty$, $-\infty < y < \infty$. This factorizes into $f_X(x).f_Y(y)$ if and only if $\rho = 0$.