

СОДЕРЖАНИЕ

Авторы.....	5
Предисловие.....	6
Введение.....	8
1. Персонализация как ведущий тренд развития нутрициологии: критерии и подходы к персонализации питания.....	12
2. Индивидуальные параметры в задачах цифровой нутрициологии.....	22
2.1. Основные биометрические параметры и их точность.....	22
2.2. Половозрастная структура населения.....	26
2.3. Распределение населения по росту и массе тела.....	32
2.4. Распределение населения по типам физической активности.....	35
3. Физиологическая потребность в энергии и пищевых веществах.....	39
3.1. Энергия.....	39
3.2. Пищевые вещества.....	41
3.3. Минорные биологически активные вещества с установленным физиологическим действием.....	45
3.4. Потребности в воде.....	47
3.5. Критически значимые пищевые вещества.....	49
3.6. Гликемический индекс пищевых продуктов.....	50

4. Структура потребления продуктов питания и нутриентов	58
4.1. Среднегодовое потребление в целом.....	58
4.2. Среднегодовое потребление отдельных домохозяйств.....	60
4.3. Изъятия из основного рациона.....	64
5. Математические модели в задачах персонализированной цифровой нутрициологии	66
5.1. Интерполяция параметров внутри класса.....	66
5.2. Индикаторы оптимальности питания.....	68
5.3. Демографическая модель.....	72
5.4. Модель баланса нутриентов.....	77
5.5. Расчет калорийности и нутриентного состава рациона.....	79
6. Базовые меню в задачах персонализированной цифровой нутрициологии	81
6.1. Составление базовых меню.....	81
6.2. Анализ сходимости величин потребленных нутриентов к рекомендованным значениям.....	94
6.3. Расчет величин порций блюд.....	98
Заключение.....	104
Список литературы.....	106

2. ИНДИВИДУАЛЬНЫЕ ПАРАМЕТРЫ В ЗАДАЧАХ ЦИФРОВОЙ НУТРИЦИОЛОГИИ

2.1. Основные биометрические параметры и их точность

С каждым пользователем связывается основное 5-мерное пространство параметров. Эти параметры следующие: 1 — пол, 2 — возраст, 3 — масса тела, 4 — рост (длина тела) или, как вариант, индекс массы тела, 5 — уровень физической активности.

Кроме перечисленных, существуют еще и другие биометрические параметры, влияющие на особенности питания, а также ряд параметров, связанных с внешними условиями. Эти параметры относятся к более сложно определяемым величинам, таким как тип метаболизма, величина подкожного жира, тип скелета. Индивидуумы различаются также своими культурными особенностями и социально-экономическим положением. На организацию структуры питания влияет экологическая обстановка в области проживания и другие внешние факторы. Создание моделей, учитывающих зависимость индивидуальных особенностей потребления продуктов питания от этих параметров, также является целью развития цифровой нутрициологии.

Кроме того, в перечень параметров, от которых зависит структура питания, входят алиментарные заболевания, которые могут быть у пользователя и с которыми связаны специфические продукты питания или их изъятие из рациона питания.

В настоящей монографии мы рассматриваем на уровне моделей только указанные выше пять основных параметров, которые имеют следующую внутреннюю классификацию для здоровых людей.

Два класса полов: мужской пол (обозначается индексом M), женский пол (обозначается индексом F).

Пять возрастных групп населения: 1 — 18–29 лет, 2 — 30–39 лет, 3 — 40–59 лет, 4 — 60 лет — 74 года, 5 — лица пожилого возраста (75 лет и старше).

Десять классов по массе тела: от 50 до 100 кг для мужчин и от 40 до 90 кг для женщин с шагом в 5 кг.

Пять ранговых параметров для классификации физической активности. Согласно методическим рекомендациям [2] для взрослого населения вводится: 1 — очень низкая активность энергозатрат, ранговый индекс $k = 1,4$; 2 — низкая активность, ранговый индекс $k = 1,6$; 3 — средняя активность, ранговый индекс $k = 1,9$; 4 — высокая активность, ранговый индекс $k = 2,2$; 5 — очень высокая активность, ранговый индекс $k = 2,5$.

Девять значений индекса I массы тела, который представляет собой отношение массы тела в килограммах к квадрату роста в метрах. Эта величина используется для оценки дефицита массы тела или избыточной массы тела. В соответствии с рекомендациями Всемирной организации здравоохранения [67] вводятся следующие градации:

- $I \leq 16$ — выраженная худоба;
- $16 < I < 16,99$ — умеренная худоба;
- $17 < I < 18,49$ — легкая худоба;
- $18,5 < I < 24,99$ — нормальные пределы;
- $I \geq 25,0$ — избыточная масса тела;
- $25,0 < I < 29,99$ — предожирение;
- $30,0 < I < 34,99$ — ожирение I степени;
- $35,0 < I < 39,99$ — ожирение II степени;
- $I \geq 40,0$ — ожирение III степени.

В то же время следует учитывать, что данная классификация отвечает среднему уровню физической активности, тогда как для высокого уровня (тяжелой физической работы или для спортсменов) этот индекс может сдвигаться вправо, а для низкой физической активности — влево.

Таким образом, в представленной укрупненной классификации размерность вектора собственных параметров пользователя, равная произведению количества независимых классов, составляет $n = 4500$ без учета внешних факторов и заболеваний. Предполагая, что каждая компонента этого вектора определяется на основе практических измерений, то есть имеет статистический характер, для оценки относительной точности ε получаемых данных по выборке конечного объема N используем формулу [65]:

$$N = -\frac{\pi n}{2\varepsilon^2} \ln(1 - (1 - \varepsilon)^2). \quad (2.1.1)$$

Эта формула получается следующим образом. Набор из n классов представляет собой гистограмму распределения N пользователей по параметрам. Предполагая, что эмпирическое распределение, которое мы обозначим как $f^{(n)}(j; N)$, где j есть номер классового интервала, приближает объективно существующую генеральную совокупность, обозначаемую $F^{(n)}(j)$, и попадание в разные классы полностью случайно, используем известный результат математической статистики о том, что статистика

$$t = \sqrt{N-1} \frac{|f^{(n)}(j; N) - F^{(n)}(j)|}{s(j; N)}$$

имеет распределение Стьюдента с $N-1$ степенью свободы. Данная статистика оценивает величину отклонения эмпирической частоты попадания в заданный классовый интервал от гипотетической генеральной вероятности. Здесь $s^2(j; N)$ есть выборочная дисперсия отдельной эмпирической частоты, равная

$$s^2(j; N) = f^{(n)}(j; N) \cdot (1 - f^{(n)}(j; N)).$$

Поскольку $N > 1$, то вместо квантиля распределения Стьюдента в оценке доверительного интервала для вероятности $f^{(n)}(j; N)$ можно взять квантиль нормального распределения $u_{1-\varepsilon/2}$, где ε — уровень значимости, на котором принимается решение о близости распределений. Тогда в приближении $N-1 \approx N$ на уровне значимости ε выражение $|f^{(n)}(j; N) - F^{(n)}(j)|$ не превосходит величины

$$s(j; N) u_{1-\varepsilon/2} / \sqrt{N}.$$

Поскольку уровень значимости не должен быть точнее интегрального уровня неопределенности в позиционировании доверительного интервала $|f^{(n)}(j; N) - F^{(n)}(j)|$, то требуется также выполнение условия:

$$\sum_{j=1}^n |f^{(n)}(j; N) - F^{(n)}(j)| \leq \varepsilon.$$

В результате получаем, что минимальное число N респондентов для достижения заданной точности ε в оценке вероятностей определяется условием:

$$\frac{u_{1-\varepsilon/2}}{\varepsilon} = \frac{\sqrt{N}}{\sum_N(n)},$$

где введено обозначение:

$$\sum_N(n) = \sum_{i=1}^n \sqrt{f^{(n)}(j; N) \cdot (1 - f^{(n)}(j; N))}.$$

Наибольшее значение последнего выражения достигается на равномерном распределении, когда $f^{(n)}(j; N) = 1/n$. Тогда $\sum_N(n) = \sqrt{n-1} \approx \sqrt{n}$, и достигается наилучшая оценка для точности ε . Функция $u_{1-\varepsilon}$ табулирована, имеет место следующая аппроксимация квантиля нормального распределения:

$$u_{1-\varepsilon/2} = \sqrt{-\frac{\pi}{2} \ln((1 - (1 - \varepsilon)^2))}.$$

В результате выписанное выше условие на минимальную длину N выборки $\frac{u_{1-\varepsilon/2}}{\varepsilon} = \frac{\sqrt{N}}{\sum_N(n)}$ приводится к виду (2.1.1).

Из формулы (2.1.1) следует, что для получения статистики с точностью $\varepsilon = 0,1$ необходимо получить данные приблизительно от $N = 1$ млн респондентов, для точности $\varepsilon = 0,05$ потребуется уже примерно 6 млн респондентов, а для точности $\varepsilon = 0,01$ — более 275 млн респондентов. Более того, если включить в число учитываемых параметров различные виды алиментарных заболеваний, то имеющаяся статистика по ним заведомо недостаточна для получения желаемой точности. Поскольку получение точности даже на уровне 10% представляется пока не вполне достижимой целью, наряду со статистическими исследованиями необходимо использовать различные теоретические модели о связи рассматриваемых параметров с калорийностью питания и с потреблением эссенциальных химических веществ и соединений.

Обратимся теперь к описанию модельных распределений населения по основным параметрам. Эти распределения получены на основе систематизации эмпирических данных.

2.2. Половозрастная структура населения

Несмотря на то, что возрастных групп введено всего пять, для анализа динамики состава этих групп по годам и изучения связи между типом питания и продолжительностью жизни требуется более детальная статистика с шагом в один год. Вообще все модели, имеющие отношение к статистике населения, основываются на демографических показателях. Эти показатели следующие: половозрастной состав населения $N^{F,M}(a, t)$ как число (в тысячах) женщин и мужчин возраста a лет в год t ; повозрастной коэффициент рождаемости $B^{F,M}(a, t)$ как число родившихся девочек и мальчиков на 1 тыс. женщин возраста a лет в год t ; повозрастной коэффициент смертности $q^{F,M}(a, t)$ как число умерших женщин и мужчин на 1 тыс. соответствующих групп возраста a лет в год t .

Хотя повозрастные коэффициенты рождаемости и смертности зависят от текущего года, то есть они не постоянны, достоверных моделей их зависимости от времени в настоящее время нет. Именно поэтому для прогноза динамики численности населения используется сценарный подход, где выбираются базовые (как в последний год), повышенные на 10–15% или аналогично пониженные значения коэффициентов. Мы будем опираться на базовый сценарий, взяв за основу данные Росстата за 2017 г. и считая коэффициенты $B^{F,M}(a, t)$ и $q^{F,M}(a, t)$ зависящими только от возраста человека. Эти данные приведены соответственно на рис. 1–3 (см. также [66]).

Пусть в год t население характеризовалось численностью $N^{F,M}(a, t)$. Тогда в следующий год $t + 1$ численность населения возраста a будет равна:

$$N^{F,M}(a, t+1) = N^{F,M}(a-1, t) \cdot (1 - q^{F,M}(a-1)). \quad (2.2.1)$$

В соответствии с формулой (2.2.1) численность возрастных классов, введенных выше, меняется со временем в зависимости

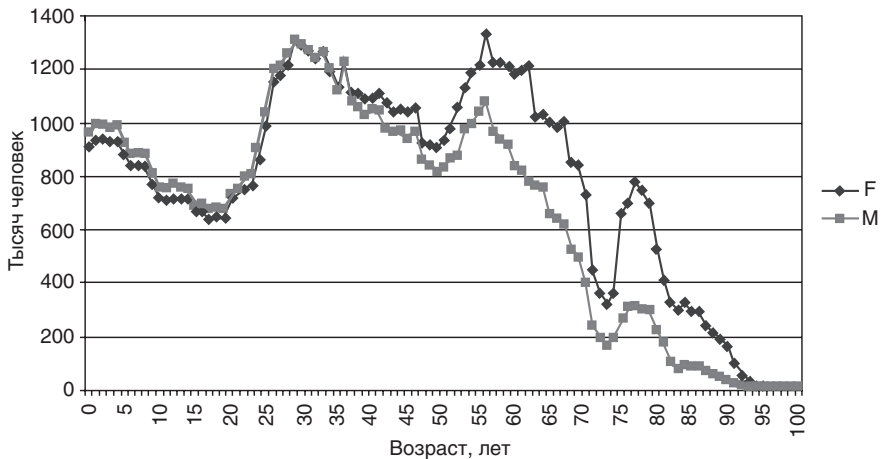


Рис. 1. Распределение населения России по полу и возрасту на 2017 г.

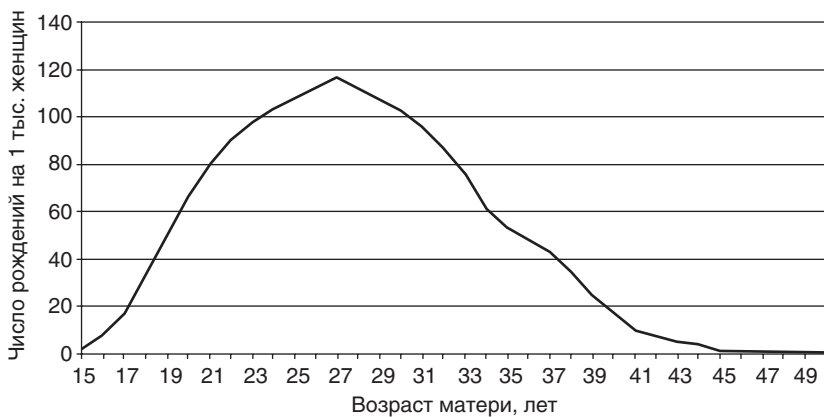


Рис. 2. Повозрастные коэффициенты рождаемости в Российской Федерации на 2016 г.

от исходной возрастной структуры населения. Например, численность первого класса в год t равна:

$$N_1^{F, M}(t) = \sum_{a=18}^{29} N^{F, M}(a, t). \quad (2.2.2)$$